(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 046 039**
A1

(12) # EUROPEAN PATENT APPLICATION

(21) Application number: 81303517.7

(22) Date of filing: 31.07.81

(51) Int. Cl.³: C 12 N 15/00
C 12 P 21/02, C 07 H 21/04
C 12 N 1/20

(30) Priority: 05.08.80 GB 8025440

(43) Date of publication of application:
17.02.82 Bulletin 82/7

(84) Designated Contracting States:
BE CH DE FR GB IT LI NL SE

(71) Applicant: G.D. Searle & Co.
Box 1045
Skokie, Illinois 60076(US)

(72) Inventor: Eaton, Michael Anthony William
43, Perry Lane Bledlow
Aylesbury Buckinghamshire(GB)

(72) Inventor: Smith, John Craig
2 Edgecote House Amersham Hill
High Wycombe Buckinghamshire(GB)

(72) Inventor: Doel, Michael Terence
Woodbury Wycombe Road
Studley Green Buckinghamshire(GB)

(72) Inventor: Lilley, David Malcolm James
"The Heven" 15 commonside Downley
High Wycombe Buckinghamshire(GB)

(72) Inventor: Carey, Norman Henry
Russells Close High Street
Chinnor Oxfordshire(GB)

(72) Inventor: Bell, Leslie David
13 Brookside
Thame Oxfordshire(GB)

(74) Representative: Smith, Sydney et al,
Elkington and Fife High Holborn House 52/54 High Holborn
London WC1V 6SH(GB)

(54) **Synthetic urogastrone gene, corresponding plasmid recombinants, transformed cells, production thereof and urogastrone expression.**

(57) A synthetic gene characterised in that it codes for the expression of urogastrone or a sub-unit thereof is disclosed.

The production thereof by the assembly and ligation of a number of nucleotide blocks is also disclosed, as are corresponding plasmid recombinants, transformed cells and the production thereof.

The expression of urogastrone is further disclosed.

Urogastrone is a polypeptide hormone (protein) which may be isolated in small amounts from human urine. It has an application in the treatment of ulcers and in the promotion of wound healing. The present invention provides a more viable commercial production.

EP 0 046 039 A1

"Synthetic urogastrone gene, corresponding plasmid recombinants, transformed cells, production thereof and urogastrone expression"

This invention relates to a synthetic urogastrone gene, to corresponding plasmid recombinants and transformed cells, to the production thereof and to urogastrone expression.

Urogastrone is a polypeptide hormone (protein) synthesised in the duodenum and in the salivary glands of normal humans, (see, for example, Heitz, et al, (1978), Gut, 19, 408-413). Urogastrone suppresses the secretion of gastric acid and promotes cell growth, (see, for example, Elder, et al, (1975), Gut, 16, 887-893). Therefore, it has an application in the treatment of ulcers and in the promotion of wound healing. Urogastrone is excreted in small amounts in human urine and may be isolated therefrom. There exists, however, a need for a more viable commercial production thereof and such is provided according to the present invention.

Urogastrone is known to consist of 53 amino acids in the following sequence:

Asn Ser Asp Ser Glu Cys Pro Leu Ser His Asp -
Gly Tyr Cys Leu His Asp Gly Val Cys Met Tyr -
Ile Glu Ala Leu Asp Lys Tyr Ala Cys Asn Cys -
Val Val Gly Tyr Ile Gly Glu Arg Cys Gln Tyr -
Arg Asp Leu Lys Trp Trp Glu Leu Arg

(see, for example, Gregory, H., and Preston, B.M., (1977), Int. J. Peptide Protein Res., 9, 107-118.)

From the above amino acid sequence, a corresponding synthetic gene sequence has been invented, subject to a number of specific non-obvious criteria, and oligonucleotide blocks synthesised which, when assembled, form a synthetic gene coding for urogastrone. The blocks have been hybridised and ligated in pre-determined stages to construct the urogastrone gene in two portions. These have been cloned in two operations into a new specifically-

designed chimeric E. coli/S. aureus vector so as to produce a full length urogastrone gene flanked only by E. coli plasmid DNA. The gene has been excised from this recombinant and re-cloned into vectors specifically designed to maximise expression of the gene in E. coli, under the control of the promoter obtained from the E. coli tryptophan operon. A protein resembling human urogastrone has thus been expressed in E. coli.

From the above amino acid sequence, because of the degeneracy of the genetic code, it is possible to predict numerous nucleotide sequences which would code for the protein.

In the inventive determination of an optimum sequence from the large number of possibilities, several non-obvious criteria have been observed. Firstly, trinucleotide codons should be used which are acceptable or preferable in the cells to be used, in particular E. coli. Secondly, it was decided that it was desirable to have different restriction enzyme recognition sites at the termini of the molecule so as to allow insertion into a plasmid in a desired orientation.. Moreover, it was decided to select sites which allowed the use of well-understood cloning vectors, such as pBR 322 (see, for example, Bolivar, F., et al, (1977), Gene, 2, 95-113). In fact, Hind III and Bam HI sites were selected and introduced at the 5' and 3' ends, respectively. Thirdly, it was thought desirable to introduce a series of restriction endonuclease recognition sites strategically placed along the molecule to enable the gene to be specifically disected to aid characterisation and, possibly mutagenesis. Also, this measure allowed the two portions of the molecule to be cloned in stages. In particular, an Xba I site was introduced at a central location in the gene. Fourthly, the synthesis should not be unnecessarily complicated and illegitimate cross-hybridisations should be minimised in order to facilitate

gene assembly. Using a computer, energies of interactions arising from all possible approximations were calculated (see, for example, Tinoco, Jr., I., et al, (1971), Nature, 230, 362-367; Powers, G.J., et al, (1975), JACS, 97, 875-889), so that stable off-diagonal interactions might be avoided where possible. Fifthly, since the protein ultimately expressed in bacterial cells will be in the form of a fusion product, it was desirable to have a means of cleaving the urogastrone portion from such a fusion product. Since urogastrone is known to be insensitive to trypsin (see, for example, Gregory and Preston, loc cit), the codons specifying the dipeptide lys-lys were introduced near the end of the gene corresponding to the urogastrone N-terminus in order to serve as a substrate for tryptic proteolysis.

The present invention relates to a synthetic gene characterised in that it codes for the expression of urogastrone or a sub-unit thereof.

The particular preferred sequence selected for the coding portion of the synthetic urogastrone gene is as follows:

```
5' A A T T C C G A T A G C G A G T G T C C T C T G -
3' T T A A G G C T A T C G C T C A C A G G A G A C -
A G T C A C G A T G G T T A C T G T C T A C A T G A C G G -
T C A G T G C T A C C A A T G A C A G A T G T A C T G C C -
C G T C T G T A T G T A T A T T G A G G C T C T A G A C A -
G C A G A C A T A C A T A T A A C T C C G A G A T C T G T -
A G T A C G C G T G T A A T T G C G T T G T T G G C T A C -
T C A T G C G C A C A T T A A C G C A A C A A C C G A T G -
A T C G G T G A G C G C T G T C A G T A T C G A G A T C T -
T A G C C A C T C G C G A C A G T C A T A G C T C T A G A -
G A A A T G G T G G G A A C T T A G A
C T T T A C C A C C C T T G A A T C T
```

One aspect of the present invention is directed accordingly. A sub-unit of such a sequence also constitutes an aspect of the present invention.

In fact, bearing in mind, inter alia, the above-mentioned considerations, for convenience, a slightly longer sequence was selected which is as follows:

Coding: A G C T T A A A A A G A A T T C C G A T A G C G A G T -
Non-Coding: A T T T T T C T T A A G G C T A T C G C T C A -
G T C C T C T G A G T C A C G A T G G T T A C T G T C T -
C A G G A G A C T C A G T G C T A C C A A T G A C A G A -
A C A T G A C G G C G T C T G T A T G T A T A T T G A G -
T G T A C T G C C G C A G A C A T A C A T A T A A C T C -
G C T C T A G A C A A G T A C G C G T G T A A T T G C G -
C G A G A T C T G T T C A T G C G C A C A T T A A C G C -
T T G T T G G C T A C A T C G G T G A G C G C T G T C A -
A A C A A C C G A T G T A G C C A C T C G C G A C A G T -
G T A T C G A G A T C T G A A T G G T G G G A A C T T -
C A T A G C T C T A G A C T T T A C C A C C C T T G A A -
A G A T A A G
T C T A T T C C T A G

The whole of this expanded sequence is shown in Figure 1 of the accompanying drawings and the restriction enzyme recognition sites are shown in Figure 2 of the accompanying drawings.

Another aspect of the present invention is directed accordingly. A sub-unit of such a sequence also constitutes an aspect of the present invention.

In the above sequences, it is to be appreciated that minor variations are possible without departing from the present inventive concept.

The present invention also relates to a process for the production of such a synthetic gene or a sub-unit thereof characterised in that it comprises the assembly and ligation of a number of oligonucleotide blocks.

It was in fact decided to synthesise a molecule having the above expanded sequence by making 23 synthetic oligonucleotide blocks as illustrated in Figure 3 of the

accompanying drawings, which will assemble by single-strand overlaps to give the complete double-stranded nucleotide sequence.

In order to minimise strong illegitimate interactions near the centre of the molecule during assembly, certain blocks of modified sequence have also been synthesised, as illustrated in Figure 4 of the accompanying drawings.

As mentioned above, the synthetic blocks selected are shown in Figure 3 of the accompanying drawings. The blocks may be constructed using known synthesis techniques (see, for example, Agarwal, et al, (1970), Nature, 227, 27-34; and Crea, et al, (1978), Proc. Natl. Acad. Sci. USA , 75, 5765-5769).

The synthetic methods will now be illustrated with reference to the synthesis of the tetradecanucleotide ApGpTpTpCpCpCpApCpCpApTpTpT.

The methods of building up oligonucleotides from smaller units by successive coupling reactions are well known (see, for example, Hsiung, et al, (1979), Nucleic Acid Research, 6, 1371-1385). The completely protected tetradecanucleotide was built up as shown in Figure 5 of the accompanying drawings (wherein, for convenience, protecting groups are not shown).

The condensation reactions indicated by arrows in Figure 5 were carried out by the following procedure exemplified in the synthesis of: DMTr A^{Bz}p G^{iBu}p-CNEt
ØCl ØCl
as shown in Figure 6 of the accompanying drawings. From 1.1 to 1.5 mmole of the 3'-phosphodiester component (I) was condensed with 1.0 mmole of the 5'-hydroxyl component (II) in anhydrous pyridine in the presence of from 3 to 4.5 mmoles of 2,4,6-triisopropyl-benzene-sulphonyl-tetrazolide. The reaction was left for 1 hour at room temperature or until chromatography on silica TLC plates eluted with 10% (v/v) methanol in chlorofrom showed that the hydroxyl component was exhausted. The reaction was

quenched with 5% (w/v) sodium bicarbonate solution and extracted with chloroform. The chloroform extract was dried and loaded onto a reverse phase chromatography column (ODS bonded to 15-25 micron silica). The fully protected dinucleotide product (III) was eluted with a solvent gradient from chloroform:methanol:water (2:6:3 v/v) to chloroform:methanol:water (2:6:0.5 v/v). The product (III) was extracted into chloroform and dried. The final isolated yield was 81%.

To proceed to further condensations, the terminal protecting group (DMTr or CNEt) was removed selectively using triethylamine in pyridine (CNEt) or a 2% (w/v) solution of benzene sulphonic acid in chloroform: methanol (DMTr) as shown in Figure 7 of the accompanying drawings.

At the completion of the synthesis, all of the protecting groups were removed by sequential treatment with 0.1 M tetraethylammonium fluoride in THF/pyridine/ water (8:1:1 v/v), ammonia and 80% acetic acid. The deprotected oligonucleotides were purified by ion-exchange HPLC and sequence analysis was carried out by the method of Wu, et al, (1976), Anal. Biochem., 74, 73-93.

The oligomeric blocks of nucleotides were hybridised and ligated (see, for example, Agarwal, et al, loc cit) in a series of steps, in order to minimise the possibilities for undesirable interactions, leading to the formation of the two portions as shown in Figure 8 of the accompanying drawings. The order of the additions in the assembly scheme was optimised for minimal incorrect ligations and in the case of especially difficult oligomeric blocks, notably 7 and 8, sub-molar quantities were used in order to remove all monomeric units before further additions were made.

In more detail:

Left-hand portion: Blocks 1 and 2 were ligated to form a dimer about the Hind III site. Blocks 4, 5 and

6 and 8, 9 and 10 were also ligated in the first round of the assembly scheme. Molar equivalents were used for all but block 8 where 0.75 molar equivalents were employed.

Block 3 was ligated with 1 + 2, block 7 (0.75 molar equivalent) with 4 to 6 and block 11 with 8 to 10. The 8 + 9 + 10 + 11 assembly has one flush end, hence some blunt-end dimerisation was observed.

1 to 3 and 4 to 7 were ligated and finally 8 to 11 were ligated to the resulting 1 to 7 species .

The dimeric 1 to 11 left-hand portion was then cleaved by Hind III (EC 3.1.23.21) and Xba I (EC 3.1.23.4) to generate the monomeric left-hand portion, with the correct cohesive termini to allow construction of recombinant plasmids.

Right-hand portion: Blocks 12 and 13 were ligated to form a dimer about the Xba I site and blocks 20, 23 and 22 similarly ligated to form a dimer about the Bam HI site. Blocks 14, 15 and 17, and 16, 18 and 19 were also ligated at this stage.

The 12, 13 dimer and the 14, 15, 17 assembly were ligated, as were the 16, 18, 19 assembly with the 20, 22, 23 dimer, where block 21 was used as a joining section.

These two species were then ligated to give an oligomeric molecule which was cleaved by Xba I and Bam HI (EC 3.1.23.6) to give the monomer 12 to 23 species having the correct cohesive termini to allow construction of recombinant plasmids.

The present invention further relates to a plasmid recombinant characterised in that it comprises a plasmid vector having inserted therein at an appropriate insertion site such a synthetic gene or a sub-unit thereof, the plasmid recombinant enabling translation in the correct phase for the mRNA corresponding to the inserted gene or sub-unit thereof and having a bacterial promoter upstream of and adjacent to the insertion site such that

the inserted gene or sub-unit thereof is under bacterial promoter control.

A particular vector, designated pLF 1, has been specifically designed to afford advantages for the purposes of the present invention. The inventive plasmid, pLF 1, is a 5K bp plasmid which may be propagated in E. coli and which may be constructed from pBR 322 and pUB 110 (see, for example, Gryczan, T.J., et al, (1978), J. Bacterial., 134, 318) by inserting the DNA sequence of the S. aureus plasmid pUB 110 between the EcoRI and Bam HI sites thereof (comprising approximately 870 bp) between the EcoRI and Bam HI sites of pBR 322, thereby replacing that region of pBR 322.

The present invention further relates to a process for the production of such a plasmid recombinant characterised in that it comprises inserting such a synthetic gene or a sub-unit thereof into an appropriate insertion site of an appropriate plasmid vector.

The following illustrates the present invention:

Phosphorylation of oligonucleotide blocks: In each case, 6 µg of oligomer and 60 µCi of [γ-^{32}P] ATP (>5000 Ci/mMol) were dried and redissolved in a final buffer concentration of 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 0.25 mM ATP, 10 mM β-mercaptoethanol and incubated with 4 units (1 unit is the amount that catalyses the production of 1 n mole of acid-insoluble ^{32}P after incubation for 30 minutes at 37°C according to Richardson, C.C.,(1972), Progress in Nucleic Acids Research, 2, 815) of T4 polynucleotide kinase (EC 2.7.1.78, Bethesda Research Labs) at 37°C for 15 minutes. The enzyme was subsequently inactivated by a 5 minute incubation at 100°C.

Ligation of oligonucleotide blocks: Except where indicated above, equimolar quantities (6 to 18 µg) of oligonucleotide blocks were incubated in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM dithiothreitol

(DTT) with 6 units (1 unit is the amount that catalyses the conversion of 1 n mole of ^{32}PPi into (α/β ^{32}P)-ATP in 20 minutes at 37°C according to Weiss, B., et al, (1968), J. Biol. Chem., 243, 4543) of T4 DNA ligase (EC 6.5.1.1, Bethesda Research Labs) at 25°C for from 3 to 16 hours. Ligated DNA was precipitated by addition of 2.5 vol absolute ethanol, collected by centrifugation and redissolved in water.

Purification of ligated species: Ligated oligonucleotide blocks were electrophoresed on 20% (w/v) polyacrylamide in 90 mM Tris-HCl, pH 8.3, 90 mM boric acid, 2.5 mM EDTA (TBE buffer), and the fragments located by autoradiography. Slices of gel containing the fragments were excised and the DNA electroeluted at 1 mA in TBE buffer onto 0.5 ml of DEAE cellulose (DE52, Whatman) for a few hours. After extensive washing by 0.1M NH_4OAc, 2 mM $Mg(OAc)_2$, 0.02% (w/v) SDS, 0.02 mM EDTA (AGEB buffer), the DNA was eluted from the DEAE cellulose by 2 ml of 1.1 M NaCl in AGEB buffer and precipitated by addition of 2.5 vol. absolute ethanol.

Construction of chimeric cloning vector pLF 1: In order to facilitate the two-stage cloning of urogastrone, it was desirable to construct a vector or vectors having Hind III, Xba I and Bam HI cleavage sites, see Figure 9 of the accompanying drawings. However, no readily available E. coli plasmids possess Xba I sites. It was surprisingly noticed that the S. aureus plasmid pUB 110, which cannot be propagated in E. coli (see, for example, Keggins, K.M., et al (1978), Proc. Natl. Acad. Sci. U.S.A., 75, 1423), contains a region of about 1 kbp of DNA bounded by sites for EcoRI and Bam HI having an approximately central Xba I site. Therefore, pUB 110 was cleaved with EcoRI (EC 3.1.4.32) and Bam HI and the DNA fragments electrophoresed on 5% (w/v) polyacrylamide. The approximately 1 kbp EcoRI/Bam HI fragment was removed by electroelution from

the excised gel slice onto DEAE cellulose, eluted by 1.1 M NaCl and ethanol precipitated. The E. coli plasmid pAT 153 (see, for example, Twigg, A.J., and Sherratt, D., (1980), Nature, 283, 216-218) was also cleaved with EcoRI and Bam HI, and the 3282 bp fragment purified by electroelution as above. These two purified DNA fragments were ligated in equimolar quantities using 10 units of T4 DNA ligase in 50 mM Tris -HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT at 15°C for 18 hours. After ethanol precipitation, the ligated DNA was dissolved in 100 µl 10 mM Tris pH 7.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$ and transformed into E. coli K12 HB 101 (genotype gal^-, lac^-, ara^-, pro^-, arg^-, str^r, rec A^-, r_k^-, M_k^-; see, for example, Boyer, H.W., and Roullard-Dussoix, D., J. Mol. Biol., 41, 459-472) using known methods (see, for example, Cohen, et al, (1972), Proc. Natl. Acad. Sci. U.S.A., 69, 2110-2114) and transformants resistant to 100 µg/ml ampicillin selected. Several transformants were analysed further by restriction enzyme cleavage of plasmid DNA, and one full length clone, designated pLF 1, selected for further use as a cloning vector. At a later stage, an additional Hind III restriction sequence was created at the EcoRI site.

Cloning of the synthetic urogastrone gene in pLF 1: The two portions of the assembled urogastrone gene were cloned in two transformation stages, as illustrated in Figure 10 of the accompanying drawings.

Right-hand portion: The longer fragment of Xba I, Bam HI-cleaved pLF 1 was purified by electroelution as above and ligated to a large excess of the right-hand portion assembly of the urogastrone gene using 10 units T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT at 15°C for 18 hours. This was transformed into E. coli as before, with selection for 100 µg/ml ampicillin. Several transformants were selected for plasmid analysis by restriction enzyme

cleavage analysis and one clone, designated pUR 1, used for further cloning. The sequence of the inserted urogastrone gene DNA was confirmed by chemical degradation analysis (see, for example, Maxam, A., and Gilbert, W., (1977), Proc. Natl. Acad. Sci. U.S.A., 74, 560-4).

In order to have a Hind III site for ligation to the 5' end of the left-hand portion of the gene, the EcoRI site was modified as follows: pUR 1 was cleaved with EcoRI and the resulting recessed ends filled using 5 units (1 unit is the amount that incorporates 10 n moles of total nucleotides into an acid-precipitable fraction in 30 minutes at 37°C using poly-d(A-T) as primer according to Richardson, C.C., et al, (1964), J. Biol. Chem., 239, 222) of DNA polymerase I (EC 2.7.7.7) in 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM β-mercapto-ethanol with 0.25 mM of each deoxynucleotide for 30 minutes at 15°C.

To the resulting flush ends, were ligated a large excess of synthetic Hind III linkers (Collaborative Research) by the known blunt end ligation procedure (see, for example, Ullrich, A., et al, (1977), Science, 196, 1313-1319). The DNA was ligated using 6 units of T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT at 15°C for 18 hours.

After Hind III restriction cleavage and purification, the full length plasmid was religated and transformed into E. coli as above, with selection for 100 µg/ml ampicillin. A transformant having a Hind III site, designated pUR 2, was selected for further cloning.

Left-hand portion: pUR 2 was cleaved with Hind III and Xba I restriction enzymes and the longer fragment purified by electroelution as above. This was ligated to a two-fold molar excess of the assembled left-hand portion of the urogastrone gene using 6 units T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT for 18 hours at 15°C. This was

transformed into E. coli K12 MRC 8 (genotype dap 103 hsd R met B1 glm 533 upp 1 dap 101 sup E thy A 103 deo rec A1) with selection for 100 µg/ml ampicillin. Several transformants were selected for plasmid analysis by restriction enzyme cleavage. One clone, designated pUR 1 was used for further characterisation and cloning. The sequence of the full urogastrone gene was confirmed by chemical degradation analysis. It should be noted that pUR 1 has no remaining S aureus DNA sequence present.

The present invention also relates to a cell, in particular an E. coli cell, characterised in that it comprises inserted therein such a synthetic gene or a sub-unit thereof or such a plasmid recombinant.

The present invention further relates to a process for the production of such a cell characterised in that it comprises inserting such a synthetic gene or a sub-unit thereof or such a plasmid recombinant into a cell.

Expression of urogastrone in E. coli: The urogastrone gene insert was cleaved from pUR 1 by Hind III and Bam HI cleavage and purified by polyacrylamide gel electrophoresis and electroelution as above. This fragment was ligated to Hind III, Bam HI-cleaved pWT 121 and pWT 221, (see, for example, Tacon, W.C.A., et al, (1980), Molec. Gen. Genet. 177, 427) and the recombinant molecules used to transform E.coli MRC 8, (see, for example, Emtage, J.S., et al, (1980), Nature, 283, 171-174), with selection for 100 µg/ml ampicillin. Transformants containing full length urogastrone genes were characterised by restriction enzyme cleavage analysis and DNA was purified by isopycnic centrifugation in caesium chloride.

Expression of urogastrone-like fusion protein was induced by growth of cells in L-broth (luria broth: 1% (w/v) bacto tryptone., 0.5% (w/v) bacto yeast extract, 0.5% (w/v) NaCl, 0.2% (w/v) glucose, 0.004% (w/v)

thymine, pH 7) containing 100 µg/ml ampicillin to an A600 nm of 0.3. Following centrifugation, the cells were washed and resuspended in M9 medium lacking tryptophan, but containing 20 µg/ml 3 β-indole acrylic acid. The cells were incubated at 37°C for 4 hours. Under these conditions maximal tryptophan promoter activity is known to occur (see Tacon, et al, loc cit), and hence expression of the urogastrone fusion protein.

The present invention also relates to a process for the production of urogastrone or a sub-unit thereof characterised in that it comprises culturing such a cell and recovering expressed protein.

As mentioned above, urogastrone has an application in the treatment of ulcers, and also in other instances where the growth promoting activity thereof would be beneficial, for example, in wound healing. Conventional administration forms may be used, the active material being used in an effective amount, for example from 0.1 to 1.0 µg/kg body weight, preferably from 0.125 to 0.5 µg/kg body weight, more preferably about 0.25 µg/kg body weight, optimally together with a conventioanl pharmaceutically-acceptable carrier, diluent or adjuvant.

CLAIMS:

1. A synthetic gene characterised in that it codes for the expression of urogastrone or a sub-unit thereof.

2. A synthetic gene is claimed in claim 1 characterised in that it comprises the following sequence or a sub-unit thereof:

Coding: A A T T C C G A T A G C G A G T G T C C T C T G -
Non-coding: T T A A G G C T A T C G C T C A C A G G A G A C -

A G T C A C G A T G G T T A C T G T C T A C A T G A C G G C -
T C A G T G C T A C C A A T G A C A G A T G T A C T G C C G -

G T C T G T A T G T A T A T T G A G G C T C T A G A C A A G -
C A G A C A T A C A T A T A A C T C C G A G A T C T G T T C -

T A C G C G T G T A A T T G C G T T G T T G G C T A C A T C -
A T G C G C A C A T T A A C G C A A C A A C C G A T G T A G -

G G T G A G C G C T G T C A G T A T C G A G A T C T G A A A -
C C A C T C G C G A C A G T C A T A G C T C T A G A C T T T -

T G G T G G G A A C T T A G A
A C C A C C C T T G A A T C T

3. A synthetic gene as claimed in claim 1 or claim 2 characterised in that it comprises the following sequence or a sub-unit thereof:

Coding: A G C T T A A A A A G A A T T C C G A T A G C G A G T
Non-coding: A T T T T T C T T A A G G C T A T C G C T C A
G T C C T C T G A G T C A C G A T G G T T A C T G T C T -
C A G G A G A C T C A G T G C T A C C A A T G A C A G A -
A C A T G A C G G C G T C T G T A T G T A T A T T G A G -
T G T A C T G C C G C A G A C A T A C A T A T A A C T C -
G C T C T A G A C A A G T A C G C G T G T A A T T G C G -
C G A G A T C T G T T C A T G C G C A C A T T A A C G C -
T T G T T G G C T A C A T C G G T G A G C G C T G T C A -
A A C A A C C G A T G T A G C C A C T C G C G A C A G T -
G T A T C G A G A T C T G A A A T G G T G G G A A C T T -
C A T A G C T C T A G A C T T T A C C A C C C T T G A A -

A G A T A A G
T C T A T T C C T A G

4. A process for the production of a synthetic gene or a sub-unit thereof as claimed in any of claims 1 to 3 characterised in that it comprises the assembly and ligation of a number of oligonucleotide blocks.

5. A plasmid recombinant characterised in that it comprises a plasmid vector having inserted therein at an appropriate insertion site a synthetic gene or a sub-unit thereof as claimed in any of claims 1 to 3, the plasmid recombinant enabling translation in the correct phase for the mRNA corresponding to the inserted gene or sub-unit thereof and having a bacterial promoter upstream of and adjacent to the insertion site such that the inserted gene or sub-unit thereof is under bacterial promoter control.

6. A plasmid recombinant as claimed in claim 5 characterised in that it comprises as the plasmid vector pLF1.

7. A process for the production of a plasmid recombinant as claimed in claim 5 characterised in that it comprises inserting a synthetic gene or a sub-unit thereof as claimed in any of claims 1 to 3 into an appropriate insertion site of an appropriate plasmid vector.

8. A cell characterised in that it comprises inserted therein a synthetic gene or a sub-unit thereof as claimed in any of claims 1 to 3 or a plasmid recombinant as claimed in claim 5 or claim 6.

9. A process for the production of a cell as claimed in claim 8 characterised in that it comprises inserting a synthetic gene or a sub-unit thereof as claimed in any of claims 1 to 3 or a plasmid recombinant as claimed in claim 5 or claim 6 into a cell.

10. A process for the production of urogastrone or a sub-unit thereof characterised in that it comprises culturing a cell as claimed in claim 8 and recovering expressed protein.

```
                                    -10                                 -5
      Met  Gln  Thr  Gln  Lys  Pro  Thr  Pro  Ser  Ser  Lys
coding A T G C A A A C A C A A A A A C C G A C T C C A A G C T C C A A G

Non-coding                                                          T T C

                                1                             5
      Leu  Lys  Lys  Asn  Ser  Asp  Ser  Glu  Cys  Pro  Leu
      C T T A A A A A G A A T T C C G A T A G C G A G T G T C C T C T G

      G A A T T T T T C T T A A G G C T A T C G C T C A C A G G A G A C

                    10                                  15
      Ser  His  Asp  Gly  Tyr  Cys  Leu  His  Asp  Gly  Val
      A G T C A C G A T G G T T A C T G T C T A C A T G A C G G C G T C

      T C A G T G C T A C C A A T G A C A G A T G T A C T G C C G C A G

            20                              25                        30
      Cys  Met  Tyr  Ile  Glu  Ala  Leu  Asp  Lys  Tyr  Ala
      T G T A T G T A T A T T G A G G C T C T A G A C A A G T A C G C G

      A C A T A C A T A T A A C T C C G A G A T C T G T T C A T G C G C
```

Fig. 1 (Part 1 of 2)

```
                                   35                                  40
Cys  Asn  Cys  Val  Val  Gly  Tyr  Ile  Gly  Glu  Arg
TGTAATTGCGTTGTTGGCTACATCGGTGAGCGC

ACATTAACGCAACAACCGATGTAGCCACTCGCG


Cys  Gln  Tyr  Arg  Asp  Leu  Lys  Trp  Trp  Glu  Leu
TGTCAGTATCGAGATCTGAAATGGTGGGAACTT

ACAGTCATAGCTCTAGACTTTACCACCCTTGAA


                  55
Arg  Ter  Gly  Ser
AGATAAGGATCC coding

TCTATTCCTAGG Non-coding
```

Fig.1 (Part 2 of 2)

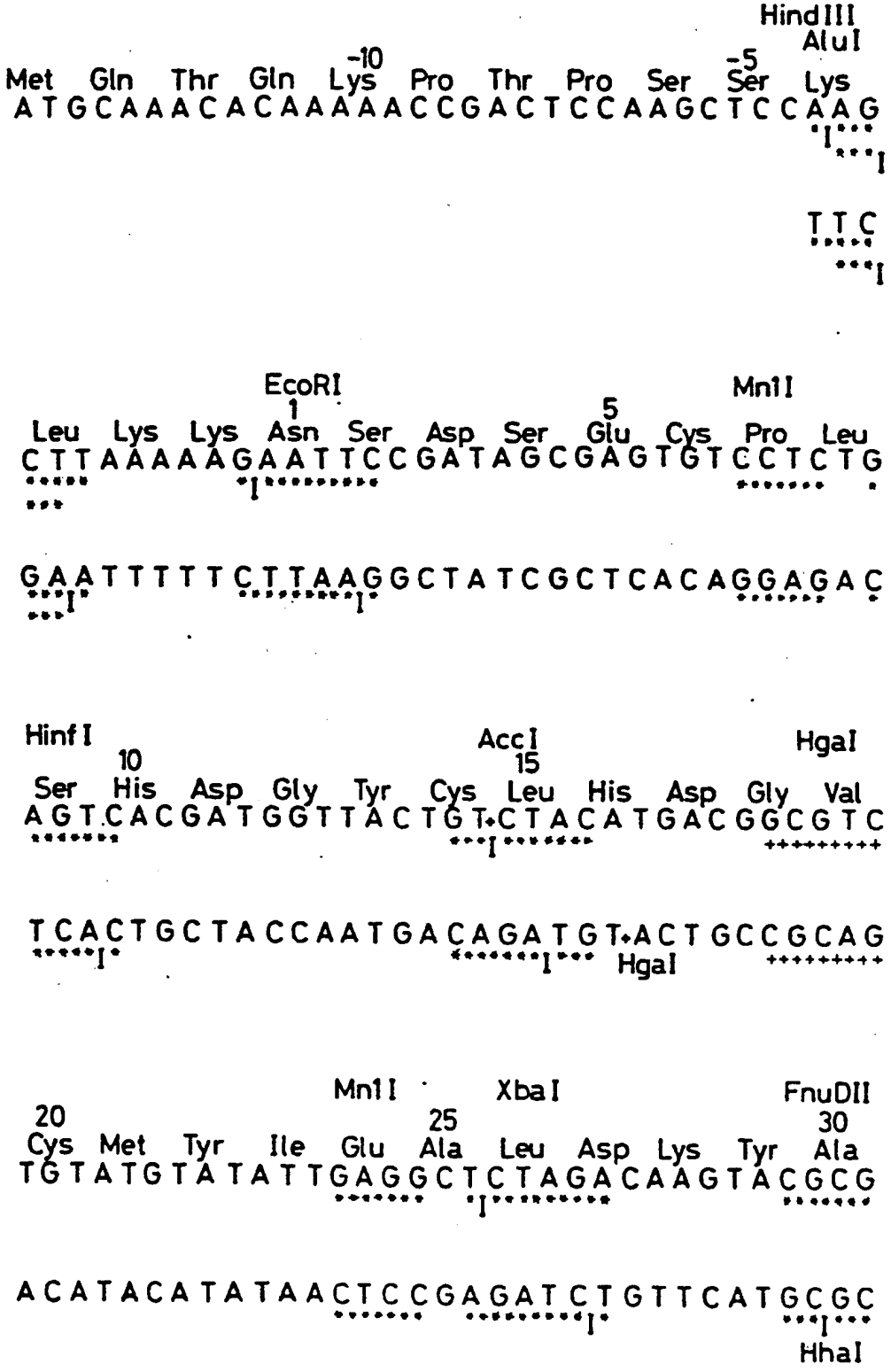


Fig. 2 (Part 1 of 2)

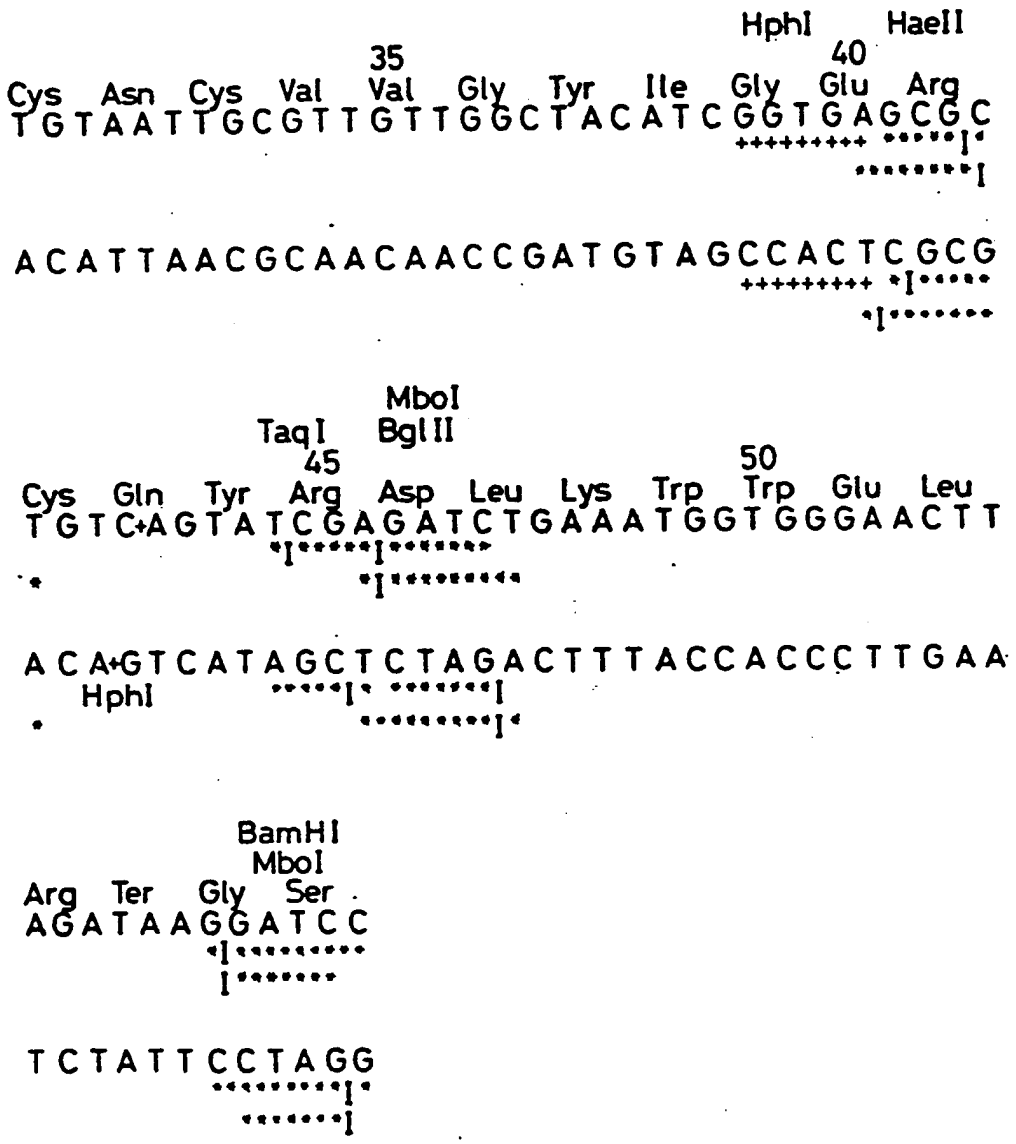
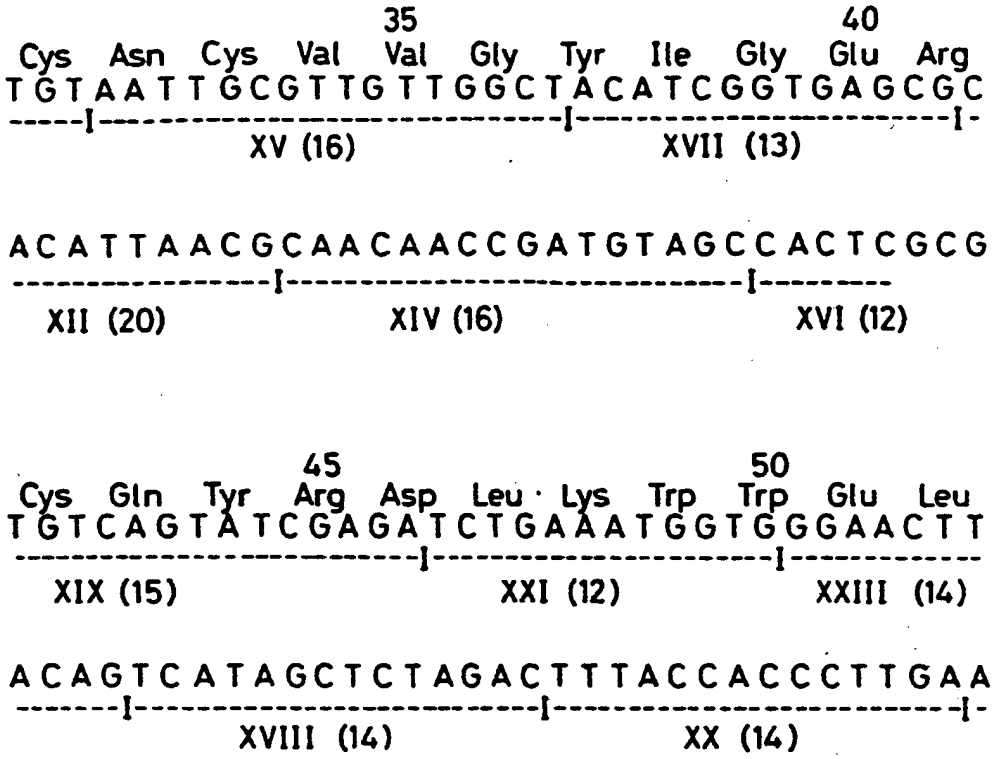


Fig. 2 (Part 2 of 2)

Fig. 3 (Part 1 of 2)



Arg Ter
AGATAAG Coding
XXIII (14)

TCTATTCCTAG Non-coding
XXII (12)

Fig.3 (Part 2 of 2)

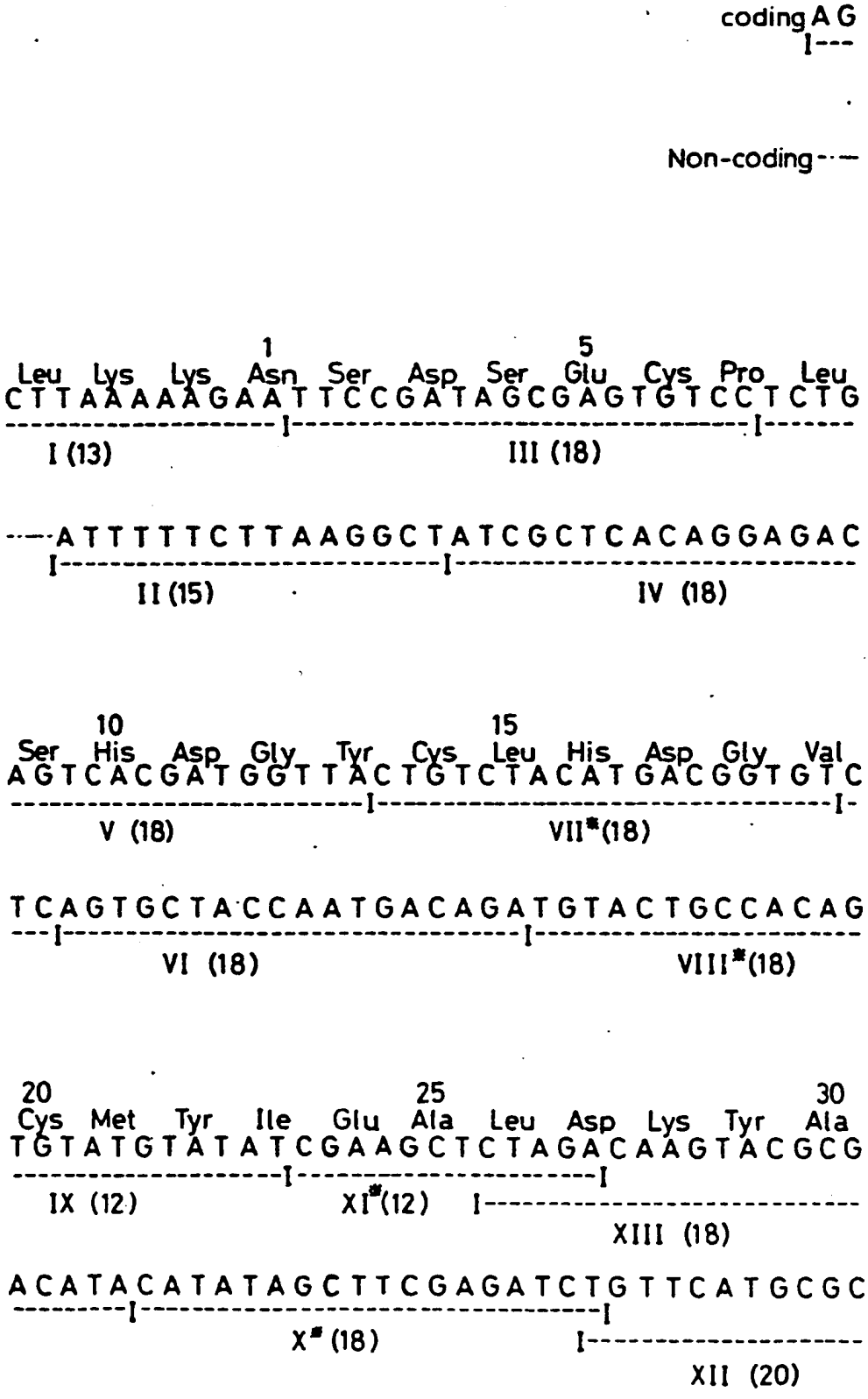


Fig. 4 (Part 1 of 2)

```
                                35                                  40
Cys  Asn  Cys  Val  Val  Gly  Tyr  Ile  Gly  Glu  Arg
TGTAATTGCGTTGTTGGCTACATCGGTGAGCGC
-----I----------------------------I--------------------I-
               XV (16)                XVII (13)

ACATTAACGCAACAACCGATGTAGCCACTCGCG
----------------I-----------------------------I---------
  XII (20)              XIV (16)                  XVI (12)

                          45                                  50
Cys  Gln  Tyr  Arg  Asp  Leu  Lys  Trp  Trp  Glu  Leu
TGTCAGTATCGAGATCTGAAATGGTGGGAACTT
----------------------I--------------------I------------
  XIX (15)               XXI (12)             XXIII (14)

ACAGTCATAGCTCTAGACTTTACCACCCTTGAA
------I----------------------I-----------------------I-
            XVIII (14)               XX (14)

Arg  Ter
AGATAAG      Coding
-----------I
XXIII (14)

TCTATTCCTAG  Non-coding
-----------
  XXII (12)
```

Fig. 4 (Part 2 of 2)

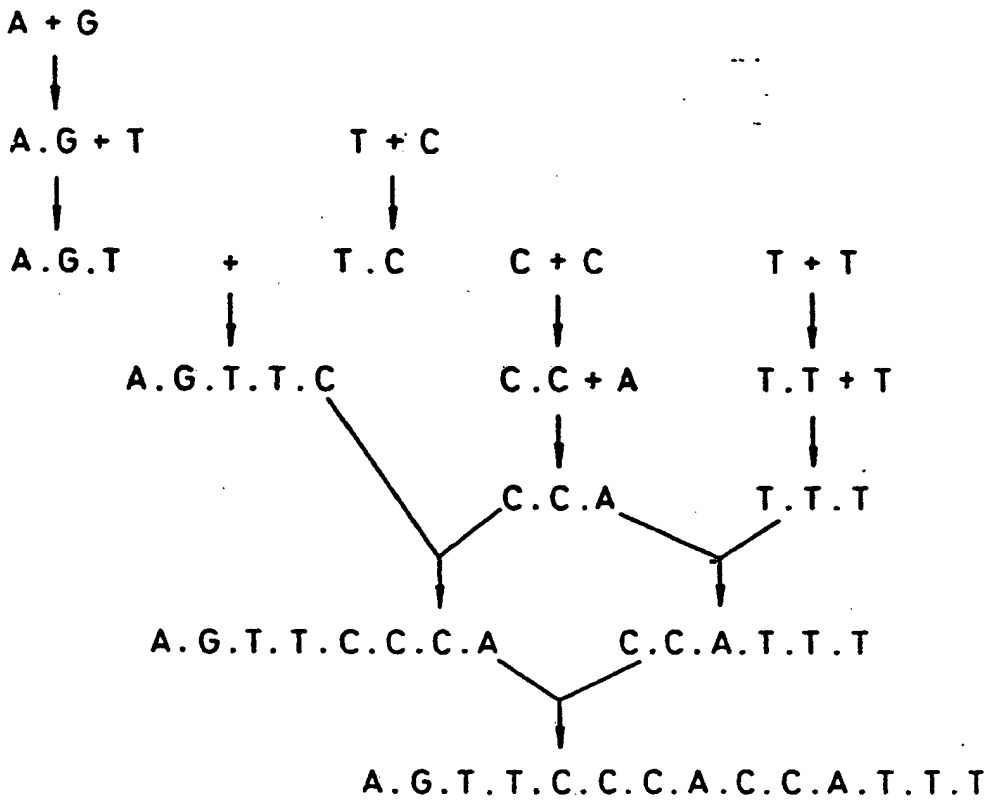

A + G
A.G + T
T + C
A.G.T
+
T.C
C + C
T + T
A.G.T.T.C
C.C + A
T.T + T
C.C.A
T.T.T
A.G.T.T.C.C.C.A
C.C.A.T.T.T
A.G.T.T.C.C.C.A.C.C.A.T.T.T

Fig.5

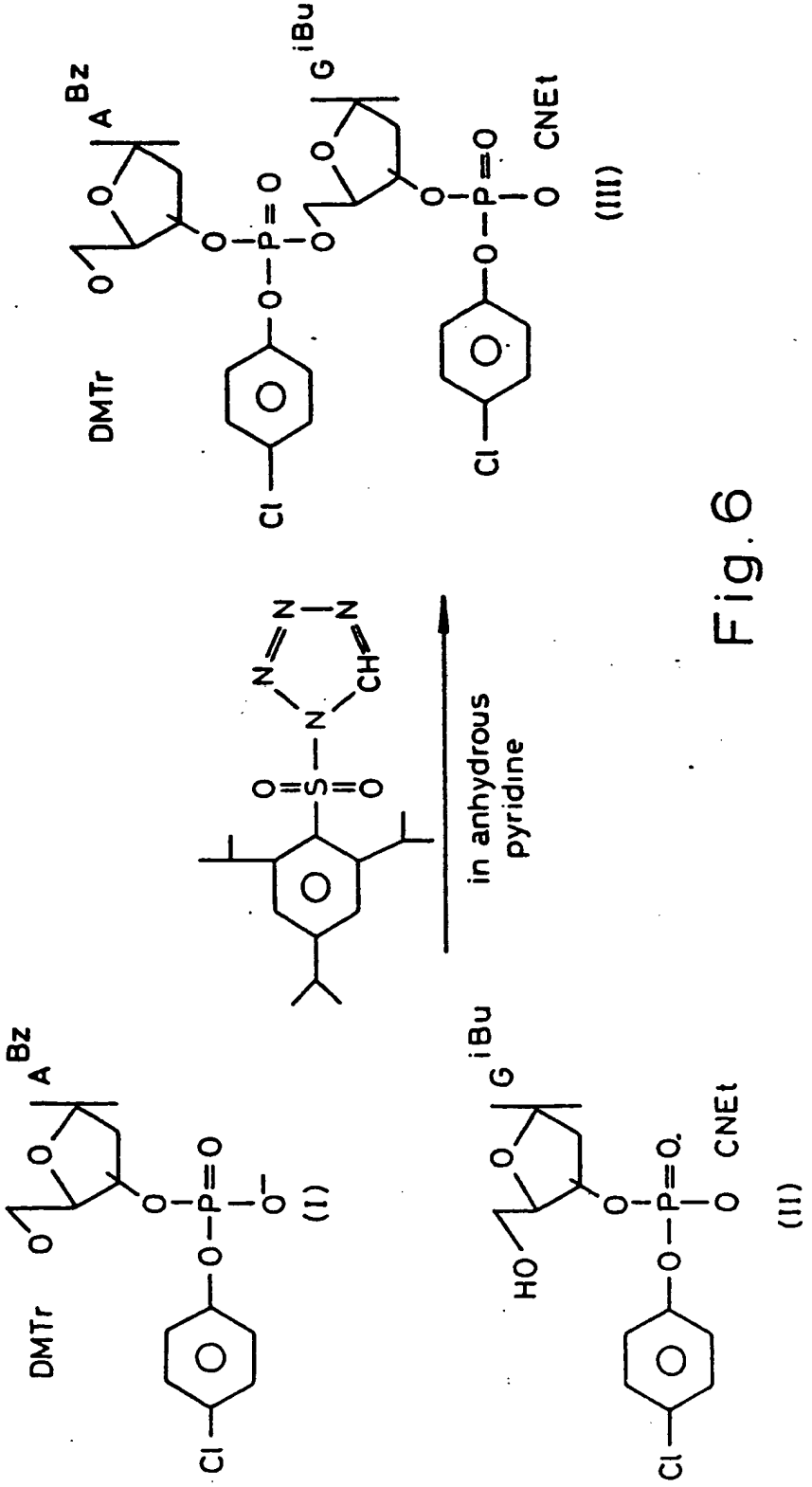


Fig. 6

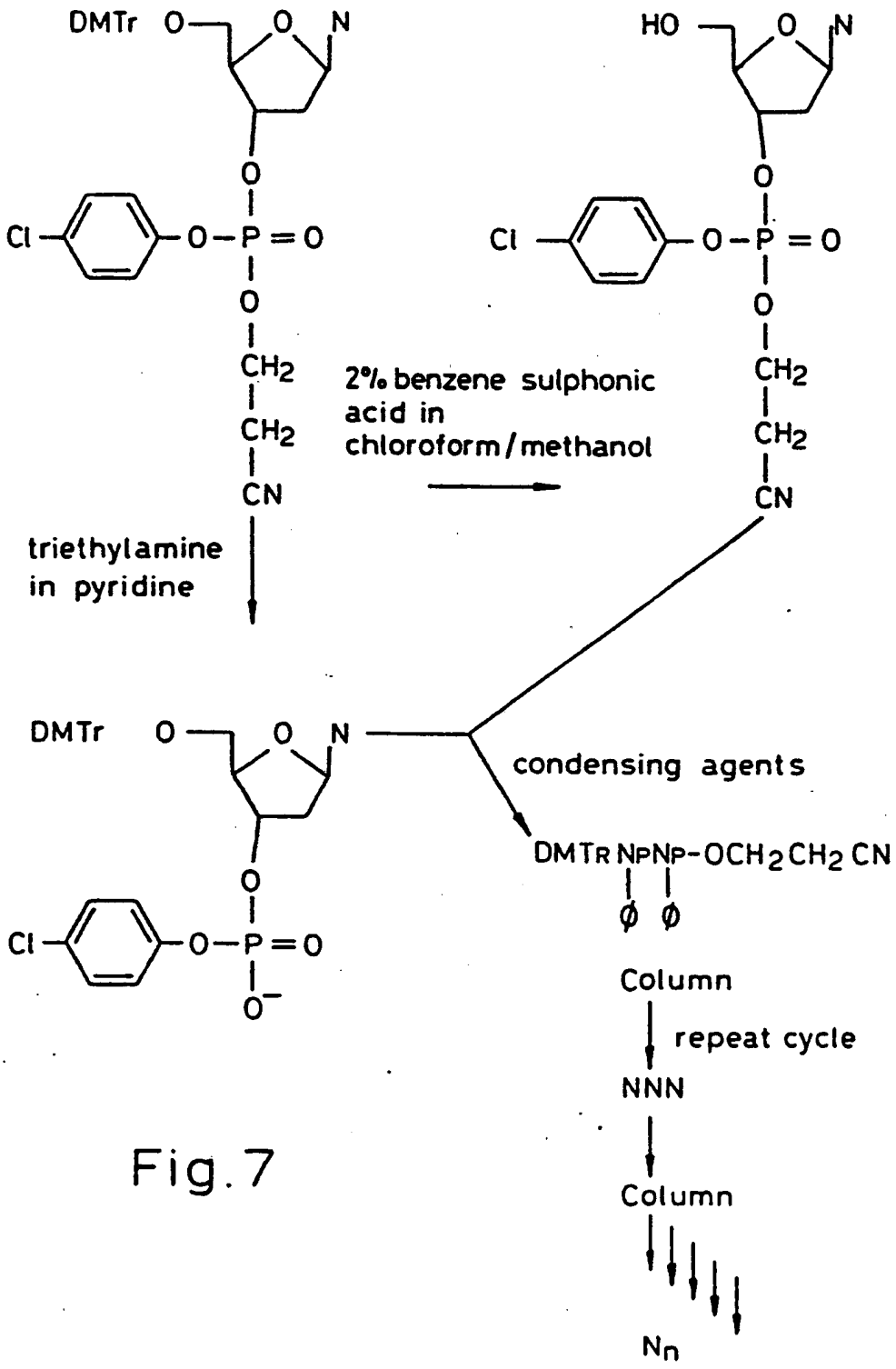

DMTr
HO
Cl
2% benzene sulphonic acid in chloroform/methanol
triethylamine in pyridine
condensing agents
DMTr NPNP-OCH2CH2CN
Column
repeat cycle
NNN
Column
Nn

Fig.7

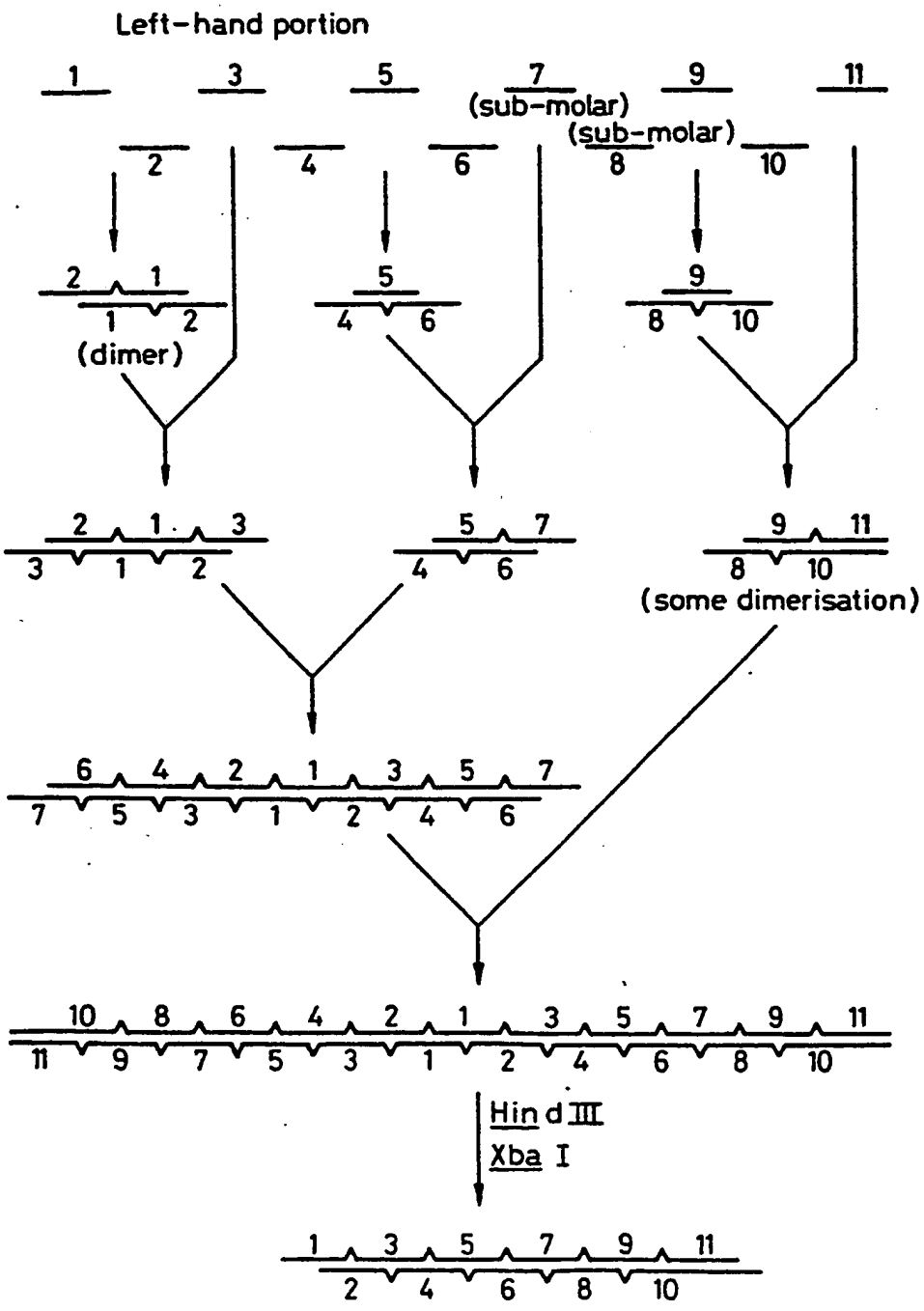


Fig.8 (Part 1 of 2)

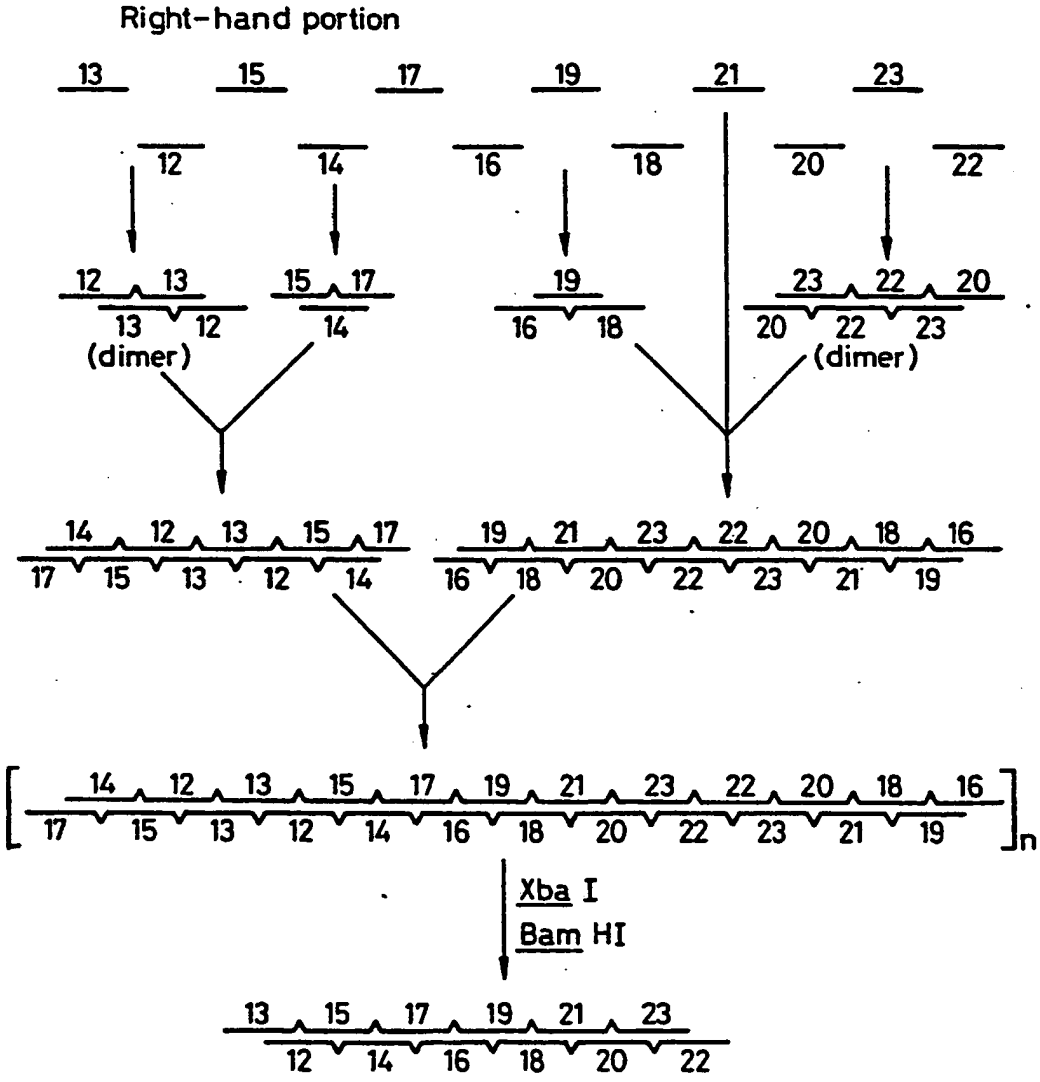


Fig.8 (Part 2 of 2)

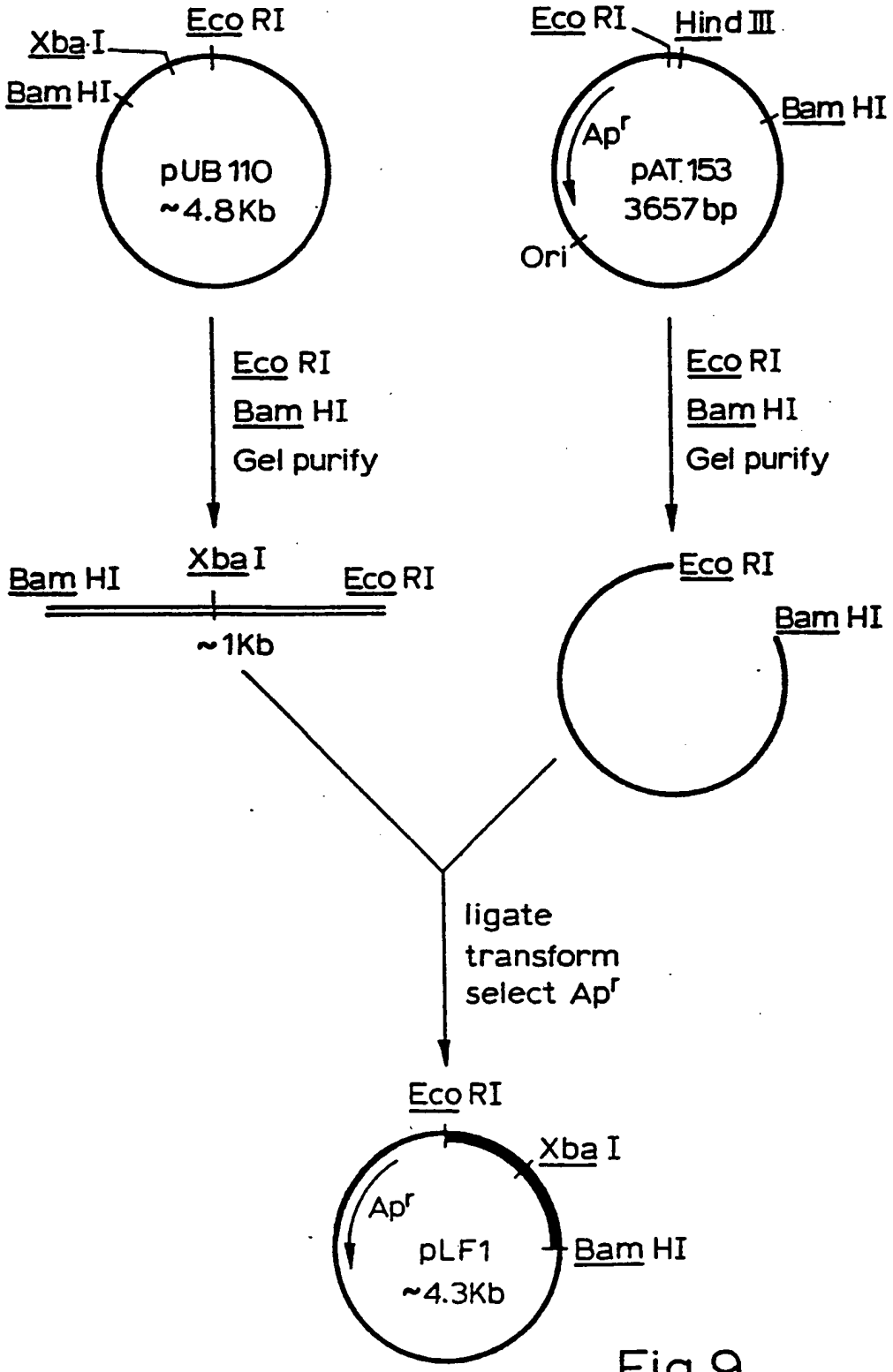

Eco RI
Xba I
Bam HI
pUB 110
~4.8Kb
Eco RI
Hind III
Bam HI
Ap^r
pAT 153
3657 bp
Ori
Eco RI
Bam HI
Gel purify
Eco RI
Bam HI
Gel purify
Bam HI
Xba I
Eco RI
~1Kb
Eco RI
Bam HI
ligate
transform
select Ap^r
Eco RI
Xba I
Ap^r
pLF1
~4.3Kb
Bam HI

Fig.9

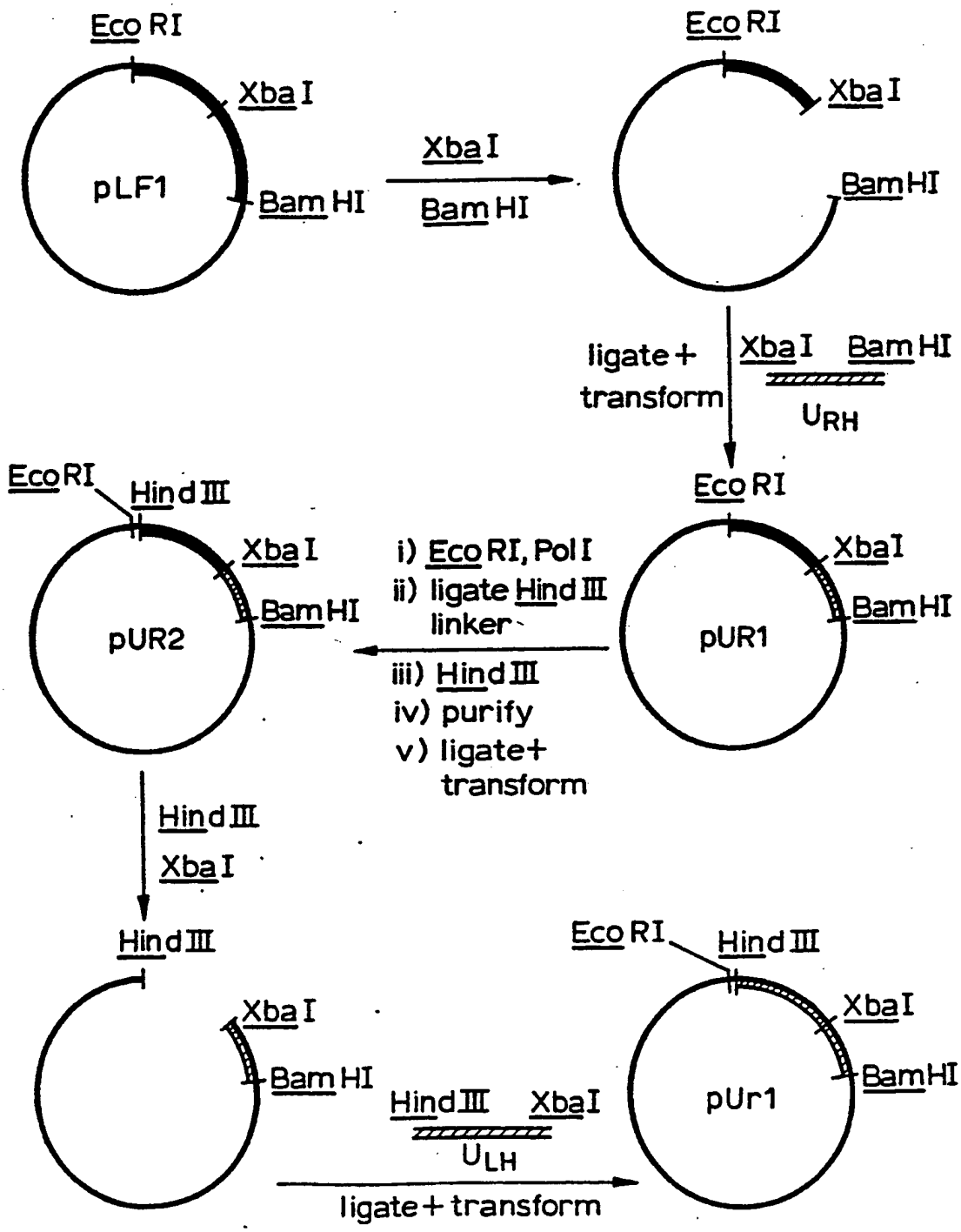

Eco RI
Xba I
pLF1
Bam HI
Xba I
Bam HI
Eco RI
Xba I
Bam HI
ligate +
transform
Xba I
Bam HI
URH
Eco RI
Hind III
Xba I
pUR2
Bam HI
i) Eco RI, Pol I
ii) ligate Hind III
linker
iii) Hind III
iv) purify
v) ligate+
transform
Eco RI
Xba I
Bam HI
pUR1
Hind III
Xba I
Hind III
Xba I
Bam HI
Hind III
Xba I
ULH
ligate + transform
Eco RI
Hind III
Xba I
Bam HI
pUr1

Fig.10

European Patent Office

EUROPEAN SEARCH REPORT

Application number EP 81 30 3517

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim
	EP - A - 0 001 931 (GENENTECH INC) * Whole document * ----	1,4,5, 8,9,10

CLASSIFICATION OF THE APPLICATION (Int. Cl.³)

C 12 N 15/00
C 12 P 21/02
C 07 H 21/04
C 12 N 1/20

TECHNICAL FIELDS SEARCHED (Int. Cl.³)

C 12 N 15/00
C 12 P 21/02
C 12 N 1/20

CATEGORY OF CITED DOCUMENTS

X: particularly relevant
A: technological background
O: non-written disclosure
P: intermediate document
T: theory or principle underlying the invention
E: conflicting application
D: document cited in the application
L: citation for other reasons

&: member of the same patent family, corresponding document

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	13-11-1981	DESCAMPS

EPO Form 1503.1 06.78

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 046 039**
B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication of patent specification: **14.01.87**

(51) Int. Cl.⁴: **C 12 N 15/00**, C 12 P 21/02, C 07 H 21/04, C 12 N 1/20

(21) Application number: **81303517.7**

(22) Date of filing: **31.07.81**

(54) Synthetic urogastrone gene, corresponding plasmid recombinants, transformed cells, production thereof and urogastrone expression.

(30) Priority: **05.08.80 GB 8025440**

(43) Date of publication of application:
17.02.82 Bulletin 82/07

(45) Publication of the grant of the patent:
14.01.87 Bulletin 87/03

(84) Designated Contracting States:
BE CH DE FR GB IT LI NL SE

(56) References cited:
EP-A-0 001 931

Dictionary of Microbiology, John Wiley and Sons, 1980, page 244

(73) Proprietor: **G.D. Searle & Co.**
P.O. Box 1045
Skokie Illinois 60076 (US)

(72) Inventor: **Eaton, Michael Anthony William**
43, Perry Lane Bledlow
Aylesbury Buckinghamshire (GB)
Inventor: **Smith, John Craig**
2 Edgecote House Amersham Hill
High Wycombe Buckinghamshire (GB)
Inventor: **Doel, Michael Terence**
Woodbury Wycombe Road
Studley Green Buckinghamshire (GB)
Inventor: **Lilley, David Malcolm James**
"The Haven" 15 commonside Downley
High Wycombe Buckinghamshire (GB)
Inventor: **Carey, Norman Henry**
Russells Close High Street
Chinnor Oxfordshire (GB)
Inventor: **Bell, Leslie David**
13 Brookside
Thame Oxfordshire (GB)

(74) Representative: **Smith, Sydney et al**
Elkington and Fife High Holborn House 52/54
High Holborn
London WC1V 6SH (GB)

Courier Press, Leamington Spa, England.

EP 0 046 039 B1

Description

This invention relates to a synthetic orogastrone gene, to corresponding plasmid recombinants and transformed cells, to the production thereof and to urogastrone expression.

Urogastrone is a polypeptide hormone (protein) synthesised in the duodenum and in the salivary glands of normal humans, (see, for example, Heitz, *et al*, (1978), Gut, *19*, 408—413). Urogastrone suppresses the secretion of gastric acid and promotes cell growth, (see, for example, Elder, *et al*, (1975), Gut, *16*, 887—893). Therefore, it has an application in the treatment of ulcers and in the promotion of wound healing. Urogastrone is excreted in small amounts in human urine and may be isolated therefrom. There exists, however, a need for a more viable commercial production thereof and such is provided according to the present invention.

Urogastrone is known to consist of 53 amino acids in the following sequence:

Asn Ser Asp Ser Glu Cys Pro Leu Ser His Asp-
Gly Tyr Cys Leu His Asp Gly Val Cys Met Tyr-
Ile Glu Ala Leu Asp Lys Tyr Ala Cys Asn Cys-
Val Val Gly Tyr Ile Gly Glu Arg Cys Gln Tyr-
Arg Asp Leu Lys Trp Trp Glu Leu Arg

(see, for example, Gregory, H., and Preston, B. M., (1977), Int. J. Peptide Protein Res., *9*, 107—118).

From the above amino acid sequence, a corresponding synthetic gene sequence has been invented, subject to a number of specific non-obvious criteria, and oligonucleotide blocks synthesised which, when assembled, form a synthetic gene coding for urogastrone. The blocks have been hybridised and ligated in pre-determined stages to construct the urogastrone gene in two portions. These have been cloned in two operations into a new specifically-designed chimeric *E. coli/S. aureus* vector so as to produce a full length urogastrone gene flanked only by *E. coli* plasmid DNA. The gene has been excised from this recombinant and re-cloned into vectors specifically designed to maximise expression of the gene in *E. coli*, under the control of the promoter obtained from the *E. coli* tryptophan operon. A protein resembling human urogastrone has thus been expressed in *E. coli*.

From the above amino acid sequence, because of the degeneracy of the genetic code, it is possible to predict numerous necleotide sequences which would code for the protein.

In the inventive determination of an optimum sequence from the large number of possibilities, several non-obvious criteria have been observed. Firstly, trinucleotide codons should be used which are acceptable or preferable in the cells to be used, in particular *E. coli*. Secondly, it was decided that it was desirable to have different restriction enzyme recognition sites at the termini of the molecule so as to allow insertion into a plasmid in a desired orientation. Moreover, it was decided to select sites which allowed the use of well-understood cloning vectors, such as pBR322 (see, for example, Bolivar, F., *et al*, (1977), Gene, *2*, 95—113). In fact, *Hind* III and *Bam* HI sites were selected and introduced at the 5' and 3' ends, respectively. Thirdly, it was thought desirable to introduce a series of restriction endonuclease recognition sites strategically placed along the molecule to enable the gene to be specifically disected to aid characterisation and, possibly mutagenesis. Also, this measure allowed the two portions of the molecule to be cloned in stages. In particular, an *Xba* I site was introduced at a central location in the gene. Fourthly, the synthesis should not be unnecessarily complicated and illegitimate cross-hybridisations should be minimised in order to facilitate gene assembly. Using a computer, energies of interactions arising from all possible approximations were calculated (see, for example, Tinoco, Jr., I., *et al*, (1971), Nature, 230, 362—367; Powers, G. J. *et al*, (1975), JACS, *97*, 875(889), so that stable off-diagonal interactions might be avoided where possible. Fifthly, since the protein ultimately expressed in bacterial cells will be in the form of a fusion product, it was desirable to have a means of cleaving the urogastrone portion from such a fusion product. Since urogastrone is known to be insensitive to trypsin (see, for example, Gregory and Preston, *loc cit*), the codons specifying the dipeptide lys-lys were introduced near the end of the gene corresponding to the urogastron N-terminus in order to serve as a substrate for tryptic proteolysis.

The present invention relates to a synthetic gene characterised in that it codes for the expression of urogastrone or an equivalent thereof or a sub-unit thereof in a bacterial cell and comprises the following sequence or sub-unit thereof:

5' A A T T C C G A T A G C G A G T G T C C T C T G-
3' T T A A G G C T A T C G C T C A C A G G A G A C-
A G T C A C G A T G G T T A C T G T C T A C A T G A C G G-
T C A G T G C T A C C A A T G A C A G A T G T A C T G C C-
C G T C T G T A T G T A T A T T G A G G C T C T A G A C A-
G C A G A C A T A C A T A T A A C T C C G A G A T C T G T-
A G T A C G C G T G T A A T T G C G T T G T T G G C T A C-
T C A T G C G C A C A T T A A C G C A A C A A C C G A T G-
A T C G G T G A G C G C T G T C A G T A T C G A G A T C T-
T A G C C A C T C G C G A C A G T C A T A G C T C T A G A-
G A A A T G G T G G G A A C T T A G A
C T T T A C C A C C C T T G A A T C T

In fact, bearing in mind, *inter alia*, the above-mentioned considerations, for convenience, a slightly longer sequence was selected which is as follows:

```
Coding 5' A G C T T A A A A A G A A T T C C G A T A G C G A G T-
Non-Coding 3' A T T T T T C T T A A G G C T A T C G C T C A-
G T C C T C T G A G T C A C G A T G G T T A C T G T C T-
C A G G A G A C T C A G T G C T A C C A A T G A C A G A-
A C A T G A C G G C G T C T G T A T G T A T A T T G A G-
T G T A C T G C C G C A G A C A T A C A T A T A A C T C-
G C T C T A G A C A A G T A C G C G T G T A A T T G C G-
C G A G A T C T G T T C A T G C G C A C A T T A A C G C-
T T G T T G G C T A C A T C G G T G A G C G C T G T C A-
A A C A A C C G A T G T A G C C A C T C G C G A C A G T-
G T A T C G A G A T C T G A A A T G G T G G G A A C T T-
C A T A G C T C T A G A C T T T A C C A C C C T T G A A-
A G A T A A G
T C T A T T C C T A G
```

The whole of this expanded sequence is shown in Figure 1 of the accompanying drawings and the restriction enzyme recognition sites are shown in Figure 2 of the accompanying drawings.

The present invention also relates to a process for the production of the above said synthetic genes or a sub-units thereof characterised in that it comprises the assembly and ligation of a number of oligonucleotide blocks.

Such "blocks" should not be regarded as sub-units of the gene. In the present context, "sub-unit" refers to a sequence which is less than the whole, but which exhibits the desired properties.

It was in fact decided to synthesise a molecule having the above expanded sequence by making 23 synthetic oligonucleotide blocks as illustrated in Figure 3 of the accompanying drawings, which will assemble by single-strand overlaps to give the complete double-stranded nucleotide sequence.

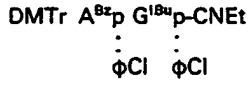
In order to minimise strong illegitimate interactions near the centre of the molecule during assembly, certain blocks of modified sequence have also been synthesised, as illustrated in Figure 4 of the accompanying drawings.

As mentioned above, the synthetic blocks selected are shown in Figure 3 of the accompanying drawings. The blocks may be constructed using known synthesis techniques (see, for example, Agarwal, *et al*, (1970), Nature, *227*, 27—34; and Crea, *et al*, (1978), Proc. Natl. Acad. Sci. U.S.A., *75*, 5765—5769).

The synthetic methods will now be illustrated with reference to the synthesis of the tetradecanucleotide ApGpTpTpCpCpCpApCpCpApTpTpT.

The methods of building up oligonucleotides from smaller units by successive coupling reactions are well known (see, for example, Hsiung, *et al*, (1979), Nucleic Acid Research, *6*, 1371—1385). The completely protected tetradecanucleotide was built up as shown in Figure 5 of the accompanying drawings (wherein, for convenience, protecting groups are not shown).

The condensation reactions indicated by arrows in Figure 5 were carried out by the following procedure exemplified in the synthesis of:



as shown in Figure 6 of the accompanying drawings. From 1.1 to 1.5 mmole of the 3'-phosphodiester component (I) was condensed with 1.0 mmole of the 5'-hydroxyl component (II) in anhydrous pyridine in the presence of from 3 to 4.5 mmoles of 2,4,6-triisopropyl-benzene-sulphonyl-tetrazolide. The reaction was left for 1 hour at room temperature or until chromatography on silica TLC plates eluted with 10% (v/v) methanol in chloroform showed that the hydroxyl component was exhausted. The reaction was quenched with 5% (w/v) sodium bicarbonate solution and extracted with chloroform. The chloroform extract was dried and loaded onto a reverse phase chromatography column (ODS bonded to 15—25 micron silica). The fully protected dinucleotide product (III) was eluted with a solvent gradient from chloroform:methanol:water (2:6:3 v/v) to chloroform:methanol:water (2:6:0.5 v/v). The product (III) was extracted into chloroform and dried. The final isolated yield was 81%.

To proceed to further condensations, the terminal protecting group (DMTr or CNEt) was removed selectively using triethylamine in pyridine (CNEt) or a 2% (w/v) solution of benzene sulphonic acid in chloroform: methanol (DMTr) as shown in Figure 7 of the accompanying drawings.

At the completion of the synthesis, all of the protecting groups were removed by sequential treatment with 0.1 M tetraethylammonium fluoride in THF/pyridine/water (8:1:1 v/v), ammonia and 80% acetic acid. The deprotected oligonucleotides were purified by ion-exchange HPLC and sequence analysis was carried out by the method of Wu, *et al*, (1976), Anal. Biochem., *74*, 73—93.

The oligomeric blocks of nucleotides were hybridised and ligated (see, for example, Agarwal, *et al*, *loc*

cit) in a series of steps, in order to minimise the possibilities for undesirable interactions, leading to the formation of the two portions as shown in Figure 8 of the accompanying drawings. The order of the additions in the assembly scheme was optimised for minimal incorrect ligations and in the case of especially difficult oligomeric blocks, notably 7 and 8, sub-molar quantities were used in order to remove all monomeric units before further additions were made.

In more detail:

Left-hand portion: Blocks 1 and 2 were ligated to form a dimer about the *Hind* III site. Blocks 4, 5 and 6 and 8, 9 and 10 were also ligated in the first round of the assembly scheme. Molar equivalents were used for all but block 8 where 0.75 molar equivalents were employed.

Block 3 was ligated with 1+2, block 7 (0.75 molar equivalent) with 4 to 6 and block 11 with 8 to 10. The 8+9+10+11 assembly has one flush end, hence some blunt-end dimerisation was observed.

1 to 3 and 4 to 7 were ligated and finally 8 to 11 were ligated to the resulting 1 to 7 species.

The dimeric 1 to 11 left-hand portion was then cleaved by *Hind* III (EC 3.1.23.21) and *Xba* I (EC 3.1.23.4) to generate the monomeric left-hand portion, with the correct cohesive termini to allow construction of recombinant plasmids.

Right-hand portion: Blocks 12 and 13 were ligated to form a dimer about the *Xba* I site and blocks 20, 23 and 22 similarly ligated to form a dimer about the *Bam* Hi site. Blocks 14, 15 and 17, and 16, 18 and 19 were also ligated at this stage.

The 12, 13 dimer and the 14, 15, 17 assembly were ligated, as were the 16, 18, 19 asesmbly with the 20, 22, 23 dimer, where block 21 was used as a joining section.

These two species were then ligated to give an oligomeric molecule which was cleaved by *Xba* I and *Bam* Hi (EC 3.1.23.6) to give the monomer 12 to 23 species having the correct cohesive termini to allow construction of recombinant plasmids.

The present invention further relates to a plasmid recombinant characterised in that it comprises a plasmid vector having inserted therein at an appropriate insertion site a synthetic gene or a sub-unit thereof as defined above, the plasmid recombinant enabling translation in the correct phase for the mRNA corresponding to the inserted gene or sub-unit thereof and having a bacterial promoter upstream of and adjacent to the insertion site such that the inserted gene or sub-unit thereof is under bacterial promoter control.

A particular vector, designated pLFI, has been specifically designed to afford advantages for the purposes of the present invention. The inventive plasmid pLFI, is a 5K bp plasmid which may be propagated in *E. coli* and which may be constructed from pBR322 and pUB110 (see, for example, Gryczan, T. J., *et al*, (1978), J. Bacterial., *134*, 318) by inserting the DNA sequence of the *S. aureus* plasmid pUB110 between the *Eco*RI and *Bam* HI sites thereof (comprising approximately 870 bp) between the *Eco*RI and *Bam* HI sites of pBR322, thereby replacing that region of pBR322.

The present invention further relates to a process for the production of such a plasmid recombinant characterised in that it comprises inserting such a synthetic gene or a sub-unit thereof as defined above into an appropriate insertion site of an appropriate plasmid vector.

The following illustrates the present invention:

Phosphorylation of oligonucleotide blocks: In each case, 6 µg of oligomer and 60 µCi of [γ-^{32}P] ATP (>5000 Ci/mMol) were dried and redissolved in a final buffer concentration of 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 0.25 mM ATP, 10 mM β-mercaptoethanol and incubated with 4 units (1 unit is the amount that catalyses the production of 1 n mole of acid-insoluble ^{32}P after incubation for 30 minutes at 37°C according to Richardson, C.C., (1972), Progress in Nucleic Acids Research, *2*, 815) of T4 polynucleotide kinase (EC 2.7.1.78, Bethesda Research Labs) at 37°C for 15 minutes. The enzyme was subsequently inactivated by a 5 minute incubation at 100°C.

Ligation of oligonucleotide blocks: Except where indicated above, equimolar quantities (6 to 18 µg) of oligonucleotide blocks were incubated in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM dithiothreitol (DTT) with 6 units (1 unit is the amount that catalyses the conversion of 1 n mole of ^{32}PPi into (α/β ^{32}P)-ATP in 20 minutes at 37°C according to Weiss, B., *et al*, (1968), J. Biol. Chem., *243* 4543) of T4 DNA ligase (EC 6.5.1.1, Bethesda Research Labs) at 25°C for from 3 to 16 hours. Ligated DNA was precipitated by addition of 2.5 vol absolute ethanol, collected by centrifugation and redissolved in water.

Purification of ligated species: Ligated oligonucleotide blocks were electrophoresed on 20% (w/v) polyacrylamide in 90 mM Tris-HCl, pH 8.3, 90 mM boric acid, 2.5 mM EDTA (TBE buffer), and the fragments located by autoradiography. Slices of gel containing the fragments were excised and the DNA electroeluted at 1 mA in TBE buffer onto 0.5 ml of DEAE cellulose (DE52, Whatman) for a few hours. After extensive washing by 0.1MNH_4OAc, 2 mM $Mg(OAc)_2$, 0.02% (w/v) SDS, 0.02 mM EDTA (AGEB buffer), the DNA was eluted from the DEAE cellulose by 2 ml of 1.1 M NaCl in AGEB buffer and precipitated by addition of 2.5 vol absolute ethanol.

Construction of chimeric cloning vector pLF1: In order to facilitate the two-stage cloning of urogastrone, it was desirable to construct a vector or vectors having *Hind* III, *Xba* I and *Bam* HI cleavage sites, see Figure 9 of the accompanying drawings. However, no readily available *E. coli* plasmids possess *Xba* I sites. It was surprisingly noticed that the *S. aureus* plasmid pUB110, which cannot be propagated in *E. coli* (see, for example, Keggins, K. M. *et al* (1978), Proc. Natl. Acad. Sci. U.S.A., *75*, 1423), contains a region of about 1kpb of DNA bounded by sites for *Eco*RI and *Bam* HI having an approximately central *Xba* I site.

Therefore, pUB110 was cleaved with *Eco*RI (EC 3.1.4.32) and *Bam* HI and the DNA fragments electrophoresed on 5% (w/v) polyacrylamide. The approximately 1kbp *Eco*RI/*Bam* HI fragment was removed by electroelution from the excised gel slice onto DEAE cellulose, eluted by 1.1 M NaCl and ethanol precipitated. The *E. coli* plasmid pAT153 (see, for example, Twigg, A. J., and Sherratt, D., (1980), Nature, *283*, 216—218) was also cleaved with *Eco*RI and *Bam* HI, and the 3282 bp fragment purified by electroelution as above. These two purified DNA fragments were ligated in equimolar quantities using 10 units of T4 DNA ligase in 50 mM Tris -HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT at 15° for 18 hours. After ethanol precipitation, the ligated DNA was dissolved in 100 µl 10 mM Tris pH 7.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$ and transformed into *E. coli* K12 HB101 (genotype gal^-, lac^-, ara^-, pro^-, arg^-, str^r, rec A^-, r_k^-, M_k^-; see, for example, Boyer, H. W., and Roullard-Dussoix, D., J. Mol. Biol., *41*, 459—472) using known methods (see, for example, Cohen, *et al*, (1972), Proc. Natl. Acad. Sci. U.S.A., *69*, 2110—2114) and transformants resistant to 100 µg/ml ampicillin selected. Several transformants were analysed further by restriction enzyme cleavage of plasmid DNA, and one full length clone, designated pLF1, selected for further use as a cloning vector. At a later stage, an additional *Hind* III restriction sequence was created at the *Eco*RI site.

Cloning of the synthetic urogastrone gene in pLF1: The two portions of the assembled urogastrone gene were cloned in two transformation stages, as illustrated in Figure 10 of the accompanying drawings.

Right-hand portion: The longer fragment of *Xba* I, *Bam* HI-cleaved pLF1 was purified by electroelution as above and ligated to a large excess of the right-hand portion assembly of the urogastrone gene using 10 units T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT at 15°C for 18 hours. This was transformed into *E. coli* as before, with selection for 100 µg/ml ampicillin. Several transformants were selected for plasmid analysis by restriction enzyme cleavage analysis and one clone, designated pUR1, used for further cloning. The sequence of the inserted urogastrone gene DNA was confirmed by chemical degradation analysis (see, for example, Maxam, A., and Gilbert, W., (1977), Proc. Natl. Acad. Sci. U.S.A., *74*, 560-4).

In order to have a *Hind* III site for ligation to the 5' end of the left-hand portion of the gene, the *Eco*RI site was modified as follows: pUR1 was cleaved with *Eco*RI and the resulting recessed ends filled using 5 units (1 unit is the amount that incorporates 10 n moles of total nucleotides into an acid-precipitable fraction in 30 minutes at 37°C using poly-d(A-T) as primer according to Richardson, C.C., *et al*, (1964), J. Biol. Chem., *239*, 222) of DNA polymerase I (EC 2.7.7.7) in 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM β-mercaptoethanol with 0.25 mM of each deoxynucleotide for 30 minutes at 15°C.

To the resulting flush ends, were ligated a large excess of synthetic *Hind* III linkers (Collaborative Research) by the known blunt end ligation procedure (see, for example, Ullrich, A., *et al*, (1977), Science, *196*, 1313—1319). The DNA was ligated using 6 units of T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATp, 20 mM DTT at 15°C for 18 hours.

After *Hind* III restriction cleavage and purification, the full length plasmid was religated and transformed into *E. coli* as above, with selection for 100 µg/ml ampicillin. A transformant having a *Hind* III site, designated pUR2, was selected for further cloning.

Left-hand portion: pUR2 was cleaved with *Hind* III and *Xba*I restriction enzymes and the longer fragment purified by electroelution as above. This was ligated to a two-fold molar excess of the assembled left-hand portion of the urogastrone gene using 6 units T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT for 18 hours at 15°C. This was transformed into *E. coli* K12 MRC 8 (genotype dap 103 hsd R met BI glm 533 upp 1 dap 101 sup E thy A 103 deo rec A1) with selection for 100 µg/ml amicillin. Several transformants were selected for plasmid analysis by restriction enzyme cleavage. One clone, designated pUR1 was used for further characterisation and cloning. The sequence of the full urogastrone gene was confirmed by chemical degradation analysis. It should be noted that pUR1 has no remaining *S aureus* DNA sequence present.

The present invention also relates to a bacterial cell, in particular an *E. coli* cell, characterised in that it comprises inserted therein a plasmid recombinant as defined above.

The present invention further relates to a process for the production of such a bacterial cell characterised in that it comprises inserting a plasmid recombinant as defined above into a bacterial cell.

Expression of urogastrone in *E. coli:* The urogastrone gene insert was cleaved from pUR1 by *Hind* III and *Bam* HI cleavage and purified by polyacrylamide gel electrophoresis and electroelution as above. This fragment was ligated to *Hind* III, *Bam* HI-cleaved pWT121 and pWT221, (see, for example, Tacon, W.C.A., *et al*, (1980), Molec. Gen. Genet. *177*, 427) and the recombinant molecules used to transform *E. coli* MRC 8, (see for example, Emtage, J. S., *et al*, (1980), Nature, *283*, 171—174), with selection for 100 µg/ml ampicillin. Transformants containing full length urogastrone genes were characterised by restriction enzyme cleavage analysis and DNA was purified by isopycnic centrifugation in caesium chloride.

Expression of urogastrone-like fusion protein was induced by growth of cells in L-broth (luria broth: 1% (w/v) bacto tryptone., 0.5% (w/v) bacto yeast extract, 0.5% (w/v) NaCl, 0.2% (w/v) glucose, 0.004% (w/v) thymine, pH 7) containing 100 µg/ml ampicillin to an A600 nm of 0.3. Following centrifugation, the cells were washed and resuspended in M9 medium lacking tryptophan, but containing 20 µgml 3 β-indole acrylic acid. The cells were incubated at 37°C for 4 hours. Under these conditions maximal tryptophan promoter activity is known to occur (see Tacon, *et al, loc cit*), and hence expression of the urogastrone fusion protein.

The present invention also relates to a process for the production of urogastrone or an equivalent

thereof or a sub-unit thereof characterised in that it comprises culturing a bacterial cell as defined above and recovering expressed protein.

The present invention further relates to a fused polypeptide characterised in that it comprises urogastrone or an equivalent thereof or a sub-unit thereof covalently bonded with all or part of a gene of the tryptophan operon, e.g. trp E, and in that it is obtained by culturing a bacterial cell as defined above and recovering expressed protein.

As mentioned above, urogastrone has an application in the treatment of ulcers, and also in other instances where the growth promoting activity thereof would be beneficial, for example, in wound healing. Conventional administration forms may be used, the active material being used in an effective amount, for example from 0.1 to 1.0 µg/kg body weight, preferably from 0.125 to 0.5 µg/kg body weight, more preferably about 0.25 µg/kg body weight, optimally together with a conventional pharmaceutically-acceptable carrier, diluent or adjuvant.

Claims

1. A synthetic gene characterised in that it codes for the expression of urogastrone or an equivalent thereof or a sub-unit thereof in a bacterial cell and comprises the following sequence or a sub-unit thereof:

Coding: 5′ A A T T C C G A T A G C G A G T G T C C T C T-
Non-coding: 3′ T T A A G G C T A T C G C T C A C A G G A G A-
G A G T C A C G A T G G T T A C T G T C T A C A T G A C G G-
C T C A G T G C T A C C A A T G A C A G A T G T A C T G C C-
C G T C T G T A T G T A T A T T G A G G C T C T A G A C A A-
G C A G A C A T A C A T A T A A C T C C G A G A T C T G T T-
G T A C G C G T G T A A T T G C G T T G T T G G C T A C A T-
C A T G C G C A C A T T A A C G C A A C A A C C G A T G T A-
C G G T G A G C G C T G T C A G T A T C G A G A T C T G A A-
G C C A C T C G C G A C A G T C A T A G C T C T A G A C T T-
A T G G T G G G A A C T T A G A
T A C C A C C C T T G A A T C T

2. A synthetic gene as claimed in Claim 1 characterised in that it comprises the following sequence or a sub-unit thereof:

Coding: 5′ A G C T T A A A A A G A A T T C C G A T A G-
Non-coding: 3′ A T T T T T C T T A A G G C T A T C-
C G A G T G T C C T C T G A G T C A C G A T G G T T A C T-
G C T C A C A G G A G A C T C A G T G C T A C C A A T G A-
G T C T A C A T G A C G G C G T C T G T A T G T A T A T T-
C A G A T G T A C T G C C G C A G A C A T A C A T A T A A-
G A G G C T C T A G A C A A G T A C G C G T G T A A T T G-
C T C C G A G A T C T G T T C A T G C G C A C A T T A A C-
C G T T G T T G G C T A C A T C G G T G A G C G C T G T C-
G C A A C A A C C G A T G T A G C C A C T C G C G A C A G-
A G T A T C G A G A T C T G A A A T G G T G G G A A C T T-
T C A T A G C T C T A G A C T T T A C C A C C C T T G A A-
A G A T A A G
T C T A T T C C T A G

3. A process for the production of a synthetic gene or a sub-unit thereof as claimed in Claim 1 or Claim 2 characterised in that it comprises the assembly and ligation of a number of oligoncleotide blocks.

4. A plasmid recombinant characterised in that it comprises a plasmid vector having inserted therein at an appropriate insertion site a synthetic gene or a sub-unit thereof as claimed in Claim 1 or Claim 2, the plasmid recombinant enabling translation in the correct phase for the mRNA corresponding to the inserted gene or sub-unit thereof and having a bacterial promoter upstream of and adjacent to the insertion site such that the inserted gene or sub-unit thereof is under bacterial promoter control.

5. A plasmid recombinant as claimed in Claim 4 characterised in that it comprises as the plasmid vector pLFI.

6. A process for the production of a plasmid recombinant as claimed in Claim 4 characterised in that it comprises inserting a synthetic gene or a sub-unit thereof as claimed in Claim 1 or Claim 2 into an appropriate insertion site of an appropriate plasmid vector.

7. A bacterial cell characterised in that it comprises inserted therein a plasmid recombinant as claimed in Claim 4 or Claim 5.

8. A process for the production of a bacterial cell as claimed in Claim 7 characterised in that it comprises inserting a plasmid recombinant as claimed in Claim 4 or Claim 5 into a bacterial cell.

9. A process for the production of urogastrone or an equivalent thereof or a sub-unit thereof characterised in that it comprises culturing a bacterial cell as claimed in Claim 7 and recovering expressed protein.

10. A fused polypeptide characterised in that it comprises urogastrone or an equivalent thereof or a sub-unit thereof covalently bonded with all or part of a gene of the tryptophan operon and in that it is obtained by culturing a bacterial cell as claimed in Claim 7 and recovering expressed protein.

11. A fused polypeptide as claimed in Claim 10 wherein the gene of the tryptophan operon is the trp E gene.

Patentansprüche

1. Synthetisches Gen, dadurch gekennzeichnet, daß es die Expression für Urogastron oder ein Äquivalent desselben oder eine Untereinheit desselben in einer Bakterienzelle codiert und die folgende Sequenz oder eine Untereinheit derselben umfaßt:

codierend: 5' A A T T C C G A T A G C G A G T G T C C T C T-
nicht codierend: 3' T T A A G G C T A T C G C T C A C A G G A G A-
G A G T C A C G A T G G T T A C T G T G T C T A C A T G A C G G-
C T C A G T G C T A C C A A T G A C A G A T G T A C T G C C-
C G T C T G T A T G T A T A T T G A G G C T C T A G A C A A-
G C A G A C A T A C A T A T A A C T C C G A G A T C T G T T-
G T A C G C G T G T A A T T G C G T T G T T G G C T A C A T-
C A T G C G C A C A T T A A C G C A A C A A C C G A T G T A-
C G G T G A G C G C T G T C A G T A T C G A G A T C T G A A-
G C C A C T C G C G A C A G T C A T A G C T C T A G A C T T-
A T G G T G G G A A C T T A G A
T A C C A C C C T T G A A T C T

2. Synthetisches Gen nach Anspruch 1, dadurch gekennzeichnet, daß es die folgende Sequenz oder eine Untereinheit derselben umfaßt:

codierend: 5' A G C T T A A A A A G A A T T C C G A T A G-
nicht codierend: 3' A T T T T T C T T A A G G C T A T C-
C G A G T G T C C T C T G A G T C A C G A T G G T T A C T-
G C T C A C A G G A G A C T C A G T G C T A C C A A T G A-
G T C T A C A T G A C G G C G T C T G T A T G T A T-
C A G A T G T A C T G C C G C A G A C A T A C A T A T A A-
G A G G C T C T A G A C A A G T A C G C G T G T A A T T G-
C T C C G A G A T C T G T T C A T G C G C A C A T T A A C-
C G T T G T T G G C T A C A T C G G T G A G C G C T G T C-
G C A A C A A C C G A T G T A G C C A C T C G C G A C A G-
A G T A T C G A G A T C T G A A A T G G T G G G A A C T T-
T C A T A G C T C T A G A C T T T A C C A C C C T T G A A-
A G A T A A G
T C T A T T C C T A G

3. Verfahren zur Erzeugung eines synthetischen Gens oder einer Untereinheit desselben nach Anspruch 1 oder Anspruch 2, dadurch gekennzeichnet, daß es den Zusammenbau und die Verknüpfung einer Anzahl von Oligonucleotid-Blöcken umfaßt.

4. Rekombinantes Plasmid, dadurch gekennzeichnet, daß es einen Plasmid-Vektor umfaßt, der ein synthetisches Gen oder eine Unterheit desselben nach Anspruch 1 oder Anspruch 2, eingefügt an einer geeigneten Insertionsstelle, aufweist, wobei das rekombinante Plasmid die Translation in der korrekten Phase für die dem eingefügten Gen oder der eingefügten Untereinheit desselben entsprechende mRNA ermöglicht und einen solchen Bakterien-Promotor von (stromaufwärts von) der Insertionsstelle und dieser benachbart aufweist, daß das eingefügte Gen oder die Untereinheit desselben der Steuerung durch den Bakterien-Promotor unterliegen.

5. Rekombinantes Plasmid nach Anspruch 4, dadurch gekennzeichnet, daß es als den Plasmid-Vektor pLFI enthält.

6. Verfahren zur Erzeugung eines rekombinanten Plasmids nach Anspruch 4, dadurch gekennzeichnet, daß es das Einfügen eines synthetischen Gens oder einer Untereinheit desselben nach Anspruch 1 oder Anspruch 2 an einer geeigneten Insertionsstelle eines geeigneten Plasmid-Vektors umfaßt.

7. Bakterienzelle, dadurch gekennzeichnet, daß sie darin eingefügt ein rekombinantes Plasmid nach Anspruch 4 oder Anspruch 5 enthält.

8. Verfahren zur Erzeugung einer Bakterienzelle nach Anspruch 7, dadurch gekennzeichnet, daß es das Einfügen eines rekombinanten Plasmids nach Anspruch 4 oder Anspruch 5 in eine Bakterienzelle umfaßt.

9. Verfahren zur Erzeugung von Urogastron oder eines Aquivalents desselben oder einer Untereinheit desselben, dadurch gekennzeichnet, daß eine Bakterienzelle nach Anspruch 7 kultiviert wird und das exprimierte Protein isoliert wird.

10. Fusioniertes Polypeptid, dadurch gekennzeichnet, daß es Urogastron oder ein Äquivalent desselben oder eine Untereinheit desselben kovalent gebunden mit dem gesamten Gen oder einem Teil des Gens des Tryptophan-Operons umfaßt und daß es durch Kultivieren einer Bakterienzelle nach Anspruch 7 und Isolierung des exprimierten Proteins erhalten wird.

11. Fusioniertes Polypeptid nach Anspruch 10, dadurch gekennzeichnet, daß das Gen des Tryptophan-Operons das trp-E-Gen ist.

Revendications

1. Gène synthétique caractérisé en ce qu'il code pour l'expression d'urogastrone ou un équivalent de cette dernière ou une sous-unité de cette dernière dans une cellule bactérienne et comprend la séquence suivante ou une sous-unité de cette dernière:

Codant: 5' A A T T C C G A T A G C G A G T G T C C T C T-
Non-codant: 3' T T A A G G C T A T C G C T C A C A G G A G A-
G A G T C A C G A T G G T T A C T G T C T A C A T G A C G G-
C T C A G T G C T A C C A A T G A C A G A T G T A C T G C C-
C G T C T G T A T G T A T A T T G A G G C T C T A G A C A A-
G C A G A C A T A C A T A T A A C T C C G A G A T C T G T T-
G T A C G C G T G T A A T T G C G T T G T T G G C T A C A T-
C A T G C G C A C A T T A A C G C A A C A A C C G A T G T A-
C G G T G A G C G C T G T C A G T A T C G A G A T C T G A A-
G C C A C T C G C G A C A G T C A T A G C T C T A G A C T T-
A T G G T G G G A A C T T A G A
T A C C A C C C T T G A A T C T

2. Gène synthétique selon la revendication 1, caractérisé en ce qu'il comprend la séquence suivante ou une sous-unité de cette dernière:

Codant: 5' A G C T T A A A A A G A A T T C C G A T A G-
Non-codant: 3' A T T T T T C T T A A G G C T A T C-
C G A G T G T C C T C T G A G T C A C G A T G G T T A C T-
G C T C A C A G G A G A C T C A G T G C T A C C A A T G A-
G T C T A C A T G A C G G C G T C T G T A T G T A T A T T-
C A G A T G T A C T G C C G C A G A C A T A C A T A T A A-
G A G G C T C T A G A C A A G T A C G C G T G T A A T T G-
C T C C G A G A T C T G T T C A T G C G C A C A T T A A C-
C G T T G T T G G C T A C A T C G G T G A G C G C T G T C-
G C A A C A A C C G A T G T A G C C A C T C G C G A C A G-
A G T A T C G A G A T C T G A A A T G G T G G G A A C T T-
T C A T A G C T C T A G A C T T T A C C A C C C T T G A A-
A G A T A A G
T C T A T T C C T A G

3. Procédé de production d'un gène synthétique ou d'une sous-unité de ce dernier selon la revendication 1 ou la revendication 2, caractérisé en ce qu'il comprend l'assemblage et la liaison d'un certain nombre de fragments d'oligonucléotides.

4. Plasmide recombinant caractérisé en ce qu'il comprend un plasmide vecteur dans lequel est inséré en un site d'insertion approprié un gène synthètique ou une sous-unité de ce dernier selon la revendication 1 ou la revendication 2, le plasmide recombinant permettant la traduction en la phase correcte pour l'ARN-m correspondant au gène inséré ou à la sous-unité de ce dernier et ayant un promoteur bactérien en amont et adjacent par rapport au site d'insertion, tel que le gène inséré ou la sous-unité de ce dernier se trouve sous le contrôle du promoteur bactérien.

5. Plasmide recombinant selon la revendication 4, caractérisé en ce qu'il comprend en tant que plasmide vecteur pLFl.

6. Procédé de production d'un plasmide recombinant selon la revendication 4, caractérisé en ce qu'il comprend l'insertion d'un gène synthétique ou d'une sous-unité de ce dernier selon la revendication 1 ou la revendication 2 dans un site d'insertion approprié d'un plasmide vecteur approprié.

7. Cellulose bactérienne caractérisée en ce que s'y trouve inséré un plasmide recombinant selon la revendication 4 ou la revendication 5.

8. Procédé de production d'une cellule bactérienne selon la revendication 7, caractérisé en ce qu'il

comprend l'insertion d'un plasmide recombinant selon la revendication 4 ou la revendication 5 dans une cellule bactérienne.

9. Procédé de production d'urogastrone ou d'un équivalent de cette dernière ou d'une sous-unité de cette dernière, caractérisé en ce qu'il comprend la culture d'une cellule bactérienne selon la revendication 7 et la récupération de la protéine exprimée.

10. Polypeptide fusionné caractérisé en ce qu'il comprend l'urogastrone ou un équivalent de cette dernière ou une sous-unité de cette dernière liée à tout ou partie d'un gène de l'opéron tryptophane et en ce qu'il est obtenu en cultivant une cellule bactérienne selon la revendication 7 et en récupérant la protéine exprimée.

11. Polypeptide fusionné selon la revendication 10, caractérisé en ce que le gène de l'opéron tryptophane est le gène trp E.

```
                                  -10                                 -5
      Met  Gln  Thr  Gln  Lys  Pro  Thr  Pro  Ser  Ser  Lys
coding A T G C A A A C A C A A A A A C C G A C T C C A A G C T C C A A G

Non-coding                                                        T T C

                        1                   5
      Leu  Lys  Lys  Asn  Ser  Asp  Ser  Glu  Cys  Pro  Leu
      C T T A A A A A G A A T T C C G A T A G C G A G T G T C C T C T G

      G A A T T T T T C T T A A G G C T A T C G C T C A C A G G A G A C

            10                        15
      Ser  His  Asp  Gly  Tyr  Cys  Leu  His  Asp  Gly  Val
      A G T C A C G A T G G T T A C T G T C T A C A T G A C G G C G T C

      T C A G T G C T A C C A A T G A C A G A T G T A C T G C C G C A G

      20                        25                        30
      Cys  Met  Tyr  Ile  Glu  Ala  Leu  Asp  Lys  Tyr  Ala
      T G T A T G T A T A T T G A G G C T C T A G A C A A G T A C G C G

      A C A T A C A T A T A A C T C C G A G A T C T G T T C A T G C G C
```

Fig. 1 (Part 1 of 2)

```
                                35                              40
Cys  Asn  Cys  Val  Val  Gly  Tyr  Ile  Gly  Glu  Arg
TGTAATTGCGTTGTTGGCTACATCGGTGAGCGC

ACATTAACGCAACAACCGATGTAGCCACTCGCG


 Cys  Gln  Tyr  Arg  Asp  Leu  Lys  Trp  Trp  Glu  Leu
TGTCAGTATCGAGATCTGAAATGGTGGGAACTT

ACAGTCATAGCTCTAGACTTTACCACCCTTGAA


           55
Arg  Ter  Gly  Ser
AGATAAGGATCC coding


TCTATTCCTAGG Non-coding
```

Fig.1 (Part 2 of 2)

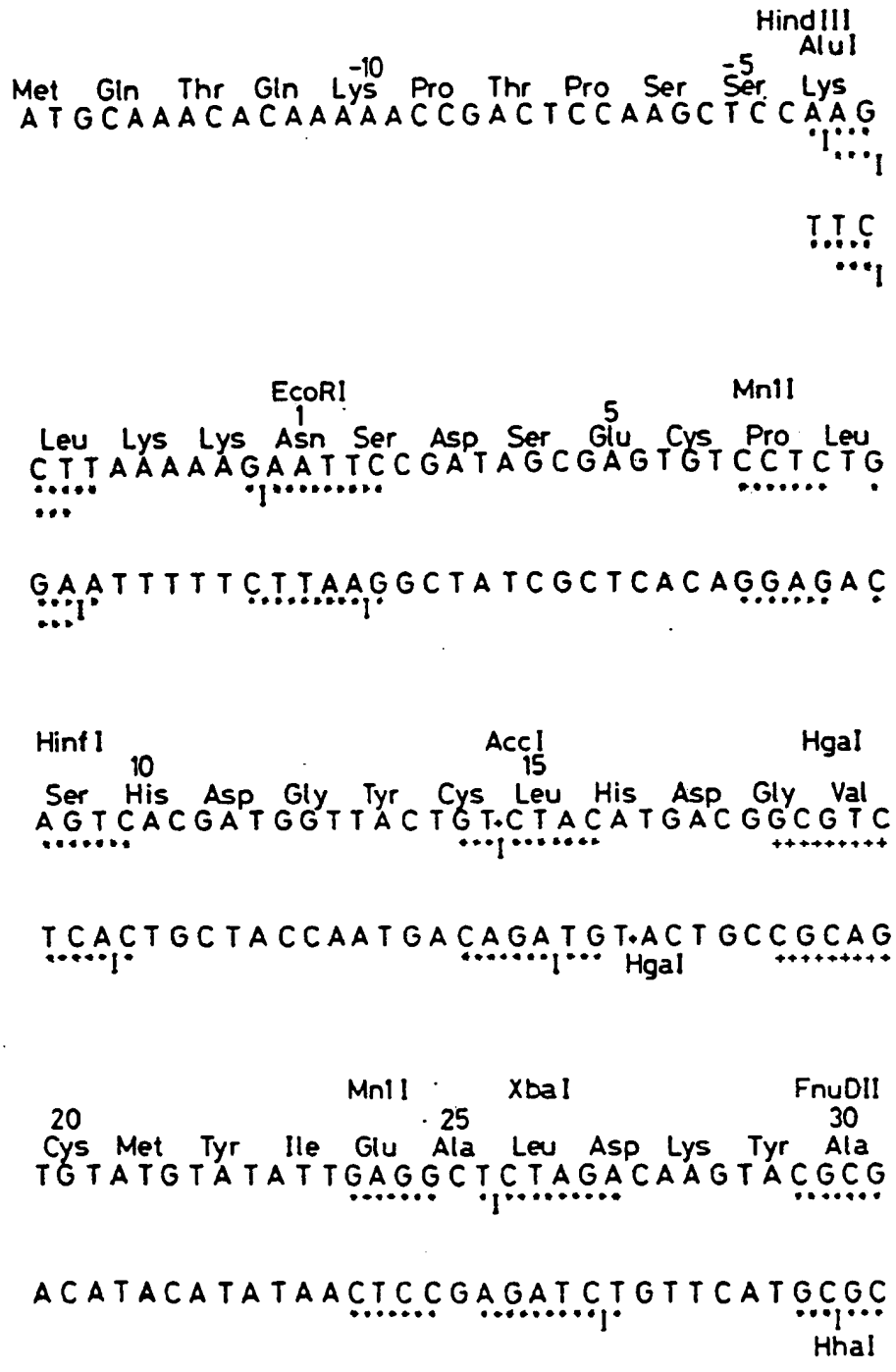


Fig. 2 (Part 1 of 2)

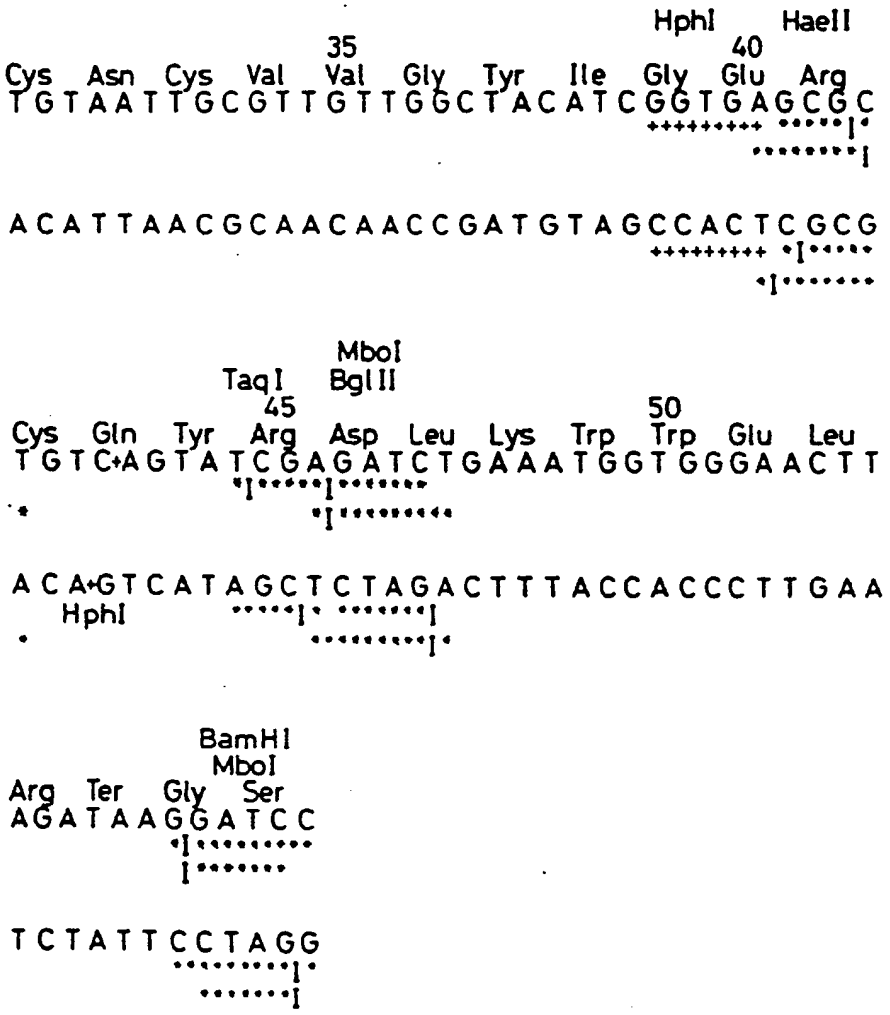


Fig. 2 (Part 2 of 2)

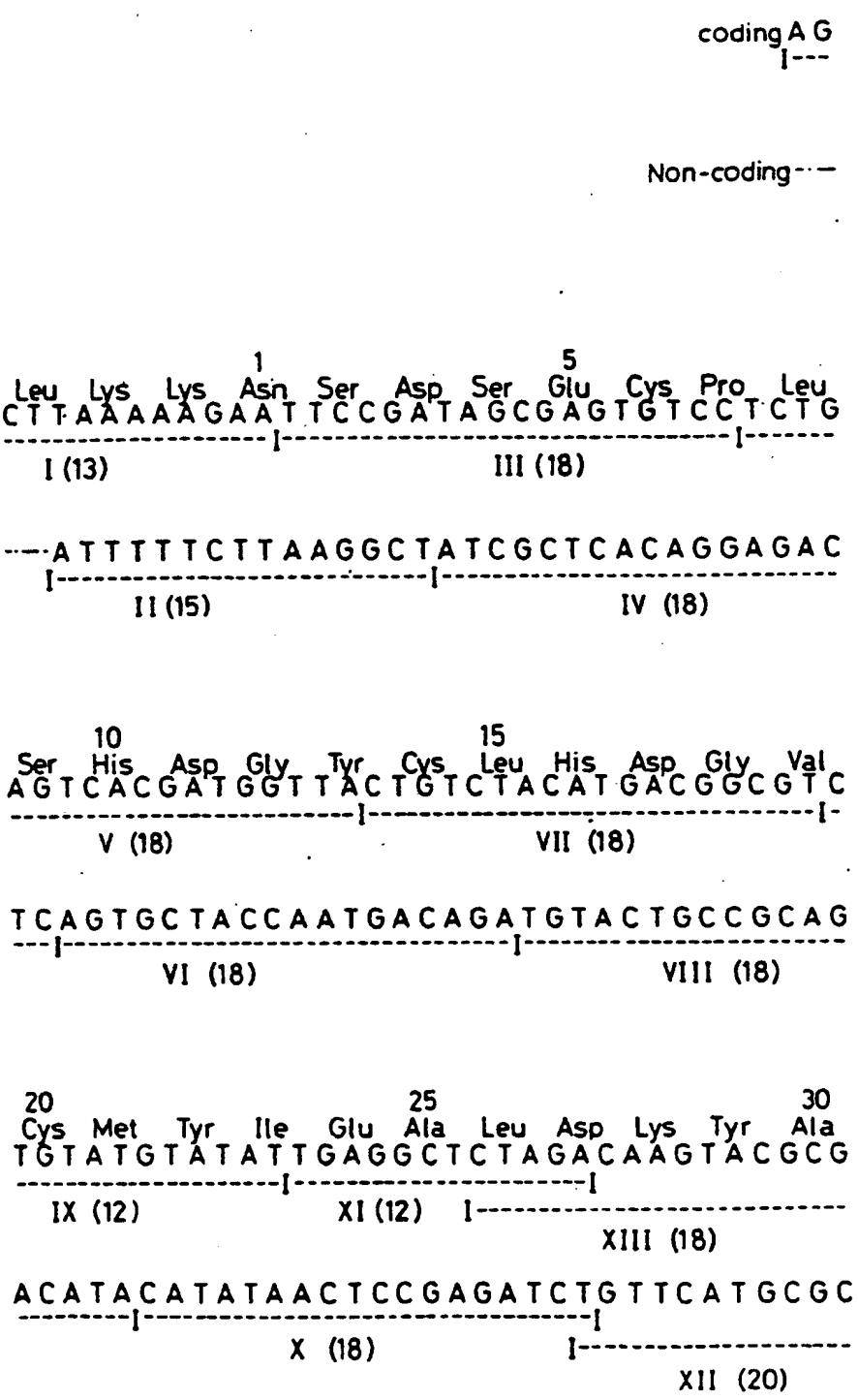


Fig. 3 (Part 1 of 2)

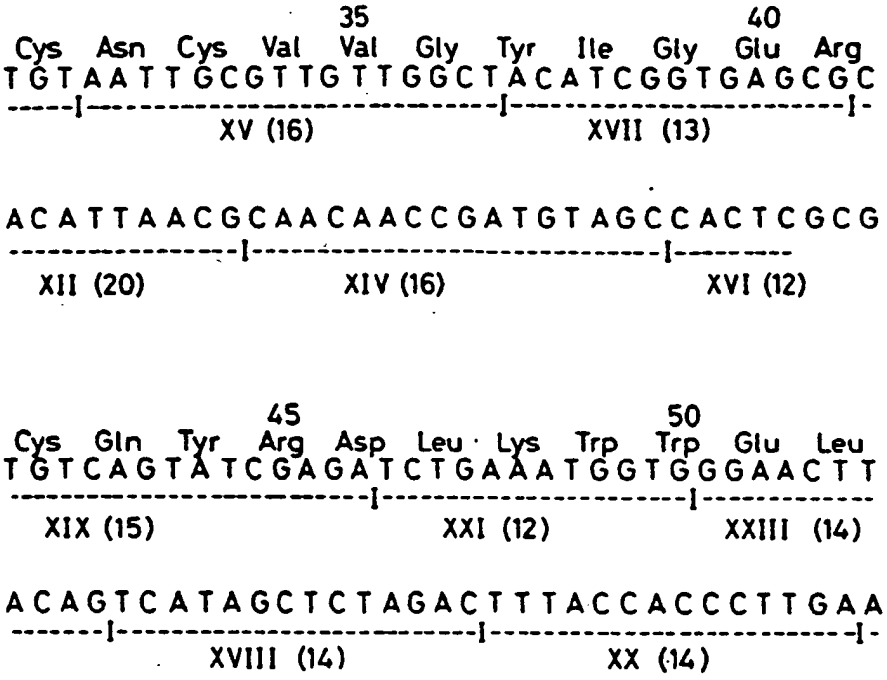


Fig.3 (Part 2 of 2)

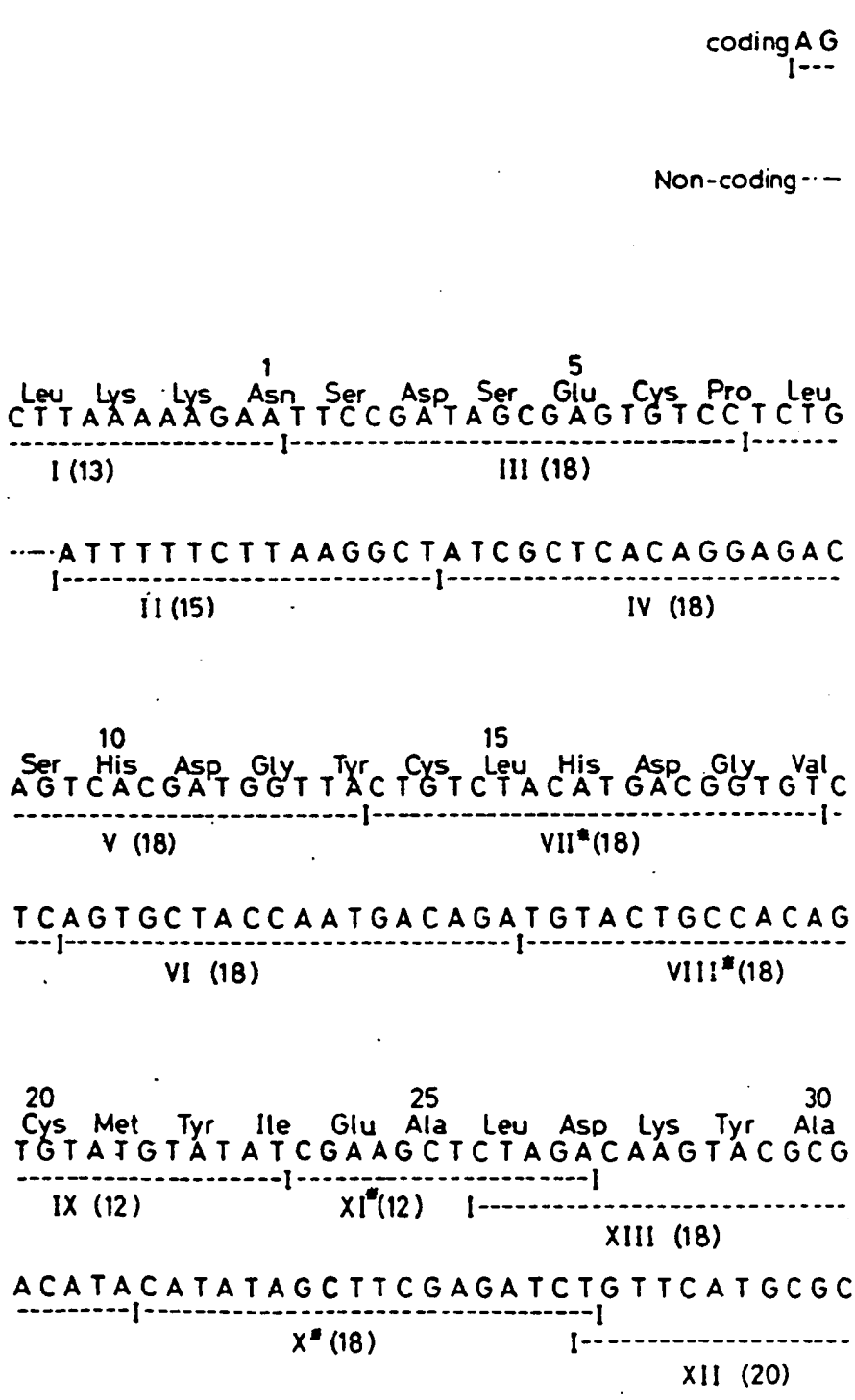


Fig. 4 (Part 1 of 2)

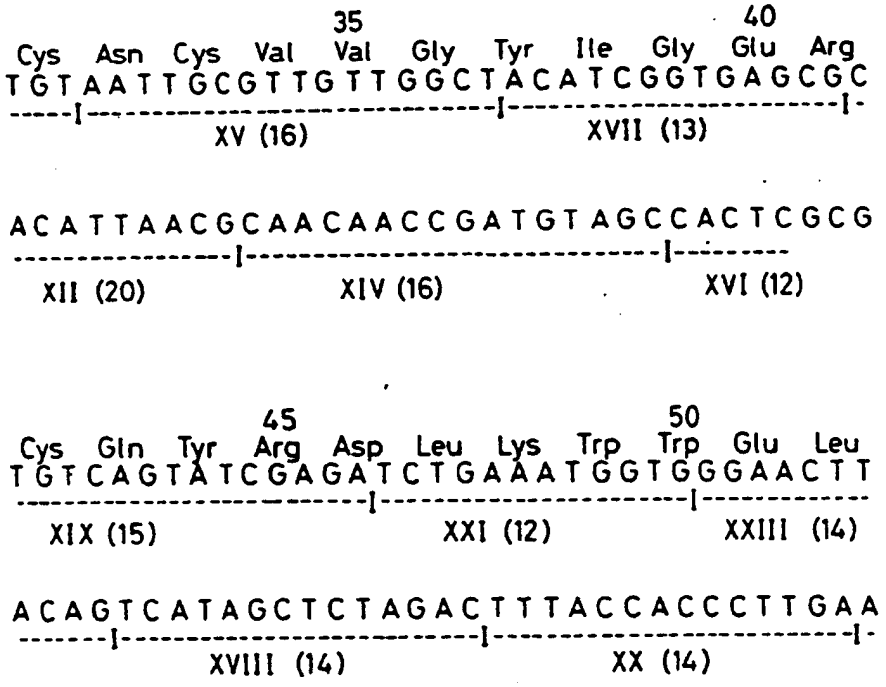


Fig. 4 (Part 2 of 2)

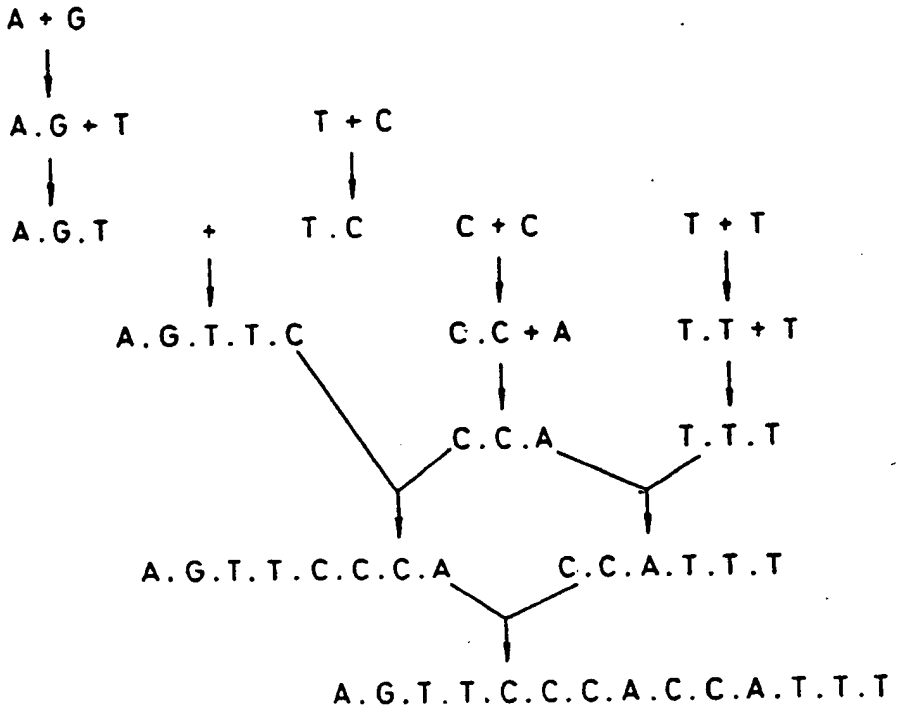

A + G
A.G + T
T + C
A.G.T
+
T.C
C + C
T + T
A.G.T.T.C
C.C + A
T.T + T
C.C.A
T.T.T
A.G.T.T.C.C.C.A
C.C.A.T.T.T
A.G.T.T.C.C.C.A.C.C.A.T.T.T

Fig.5

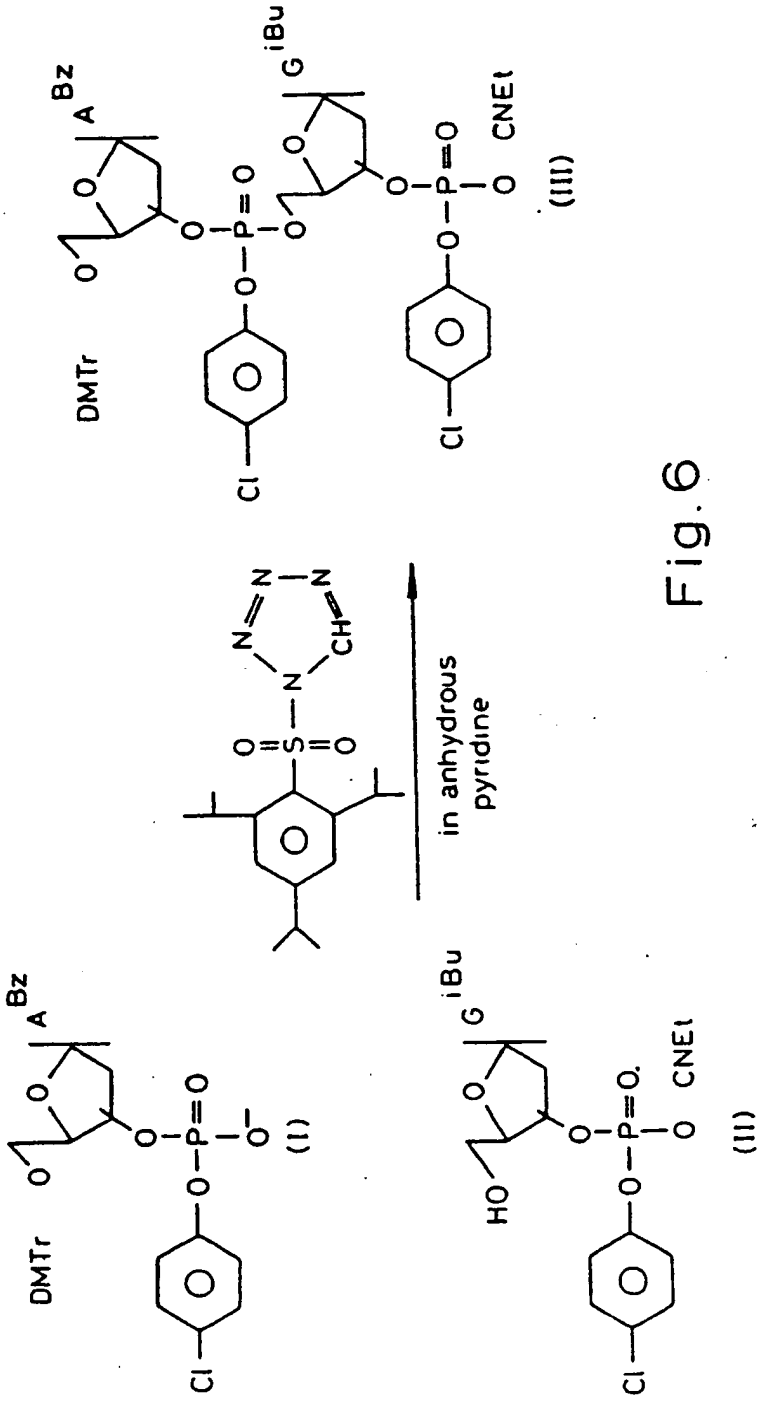


Fig. 6

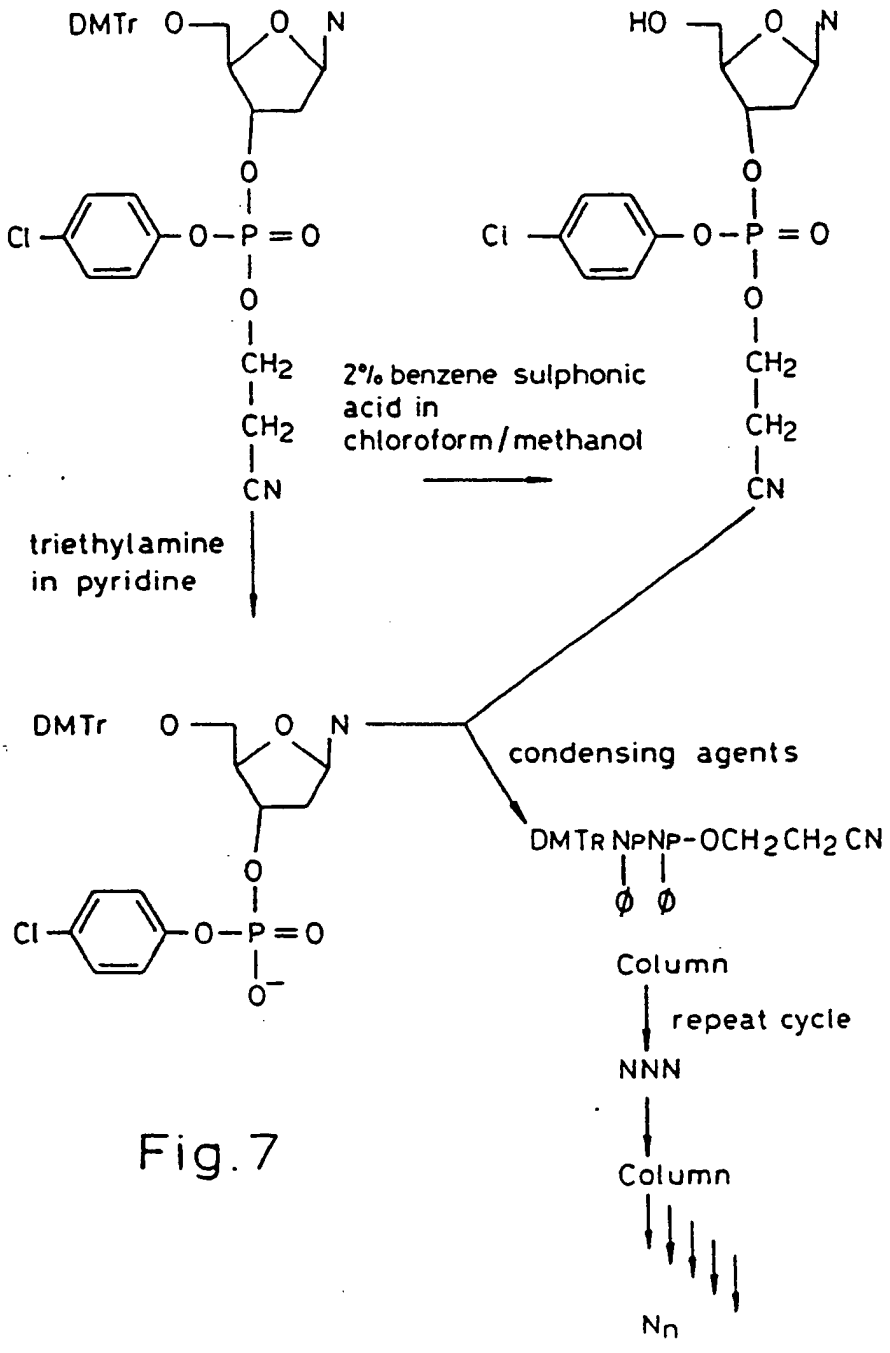

DMTr
HO
Cl
2% benzene sulphonic acid in chloroform/methanol
triethylamine in pyridine
condensing agents
DMTR NPNP-OCH2CH2 CN
Column
repeat cycle
NNN
Column
Nn

Fig.7

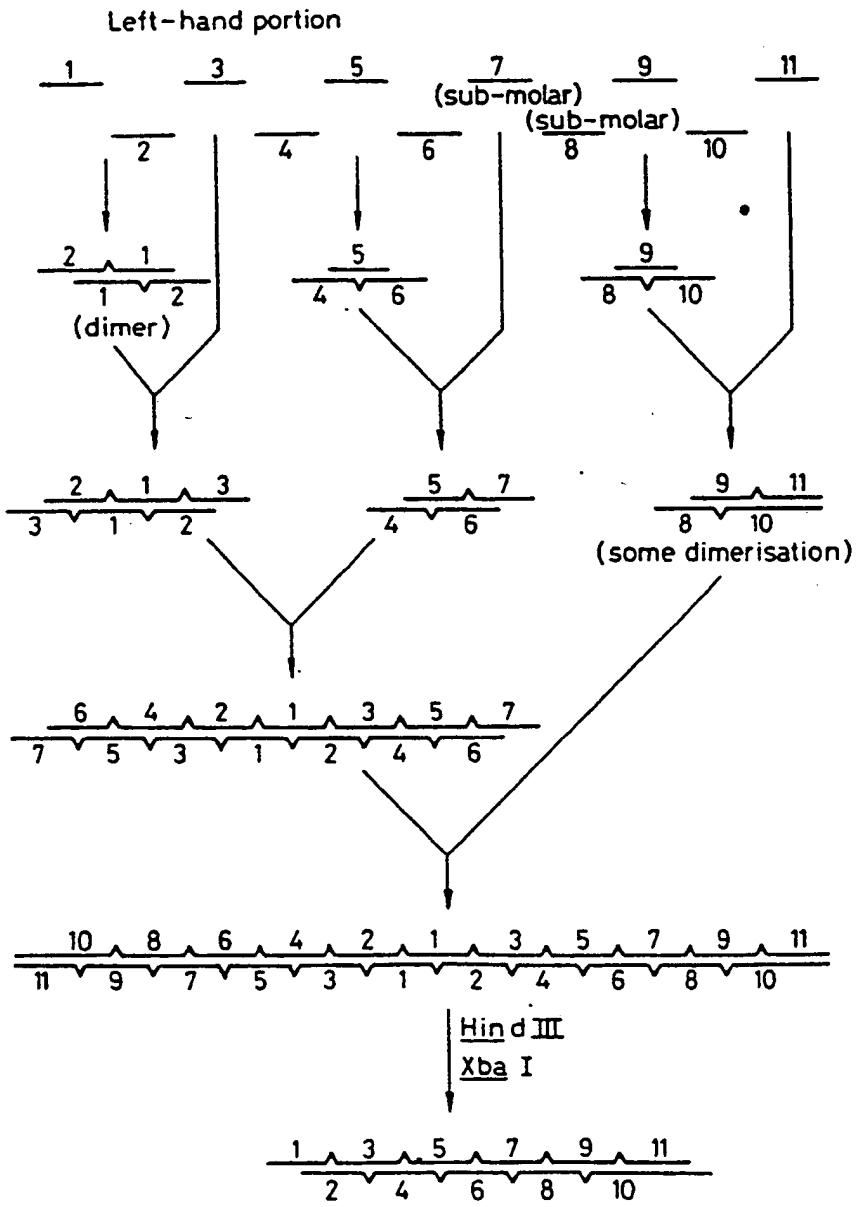


Fig.8 (Part 1 of 2)

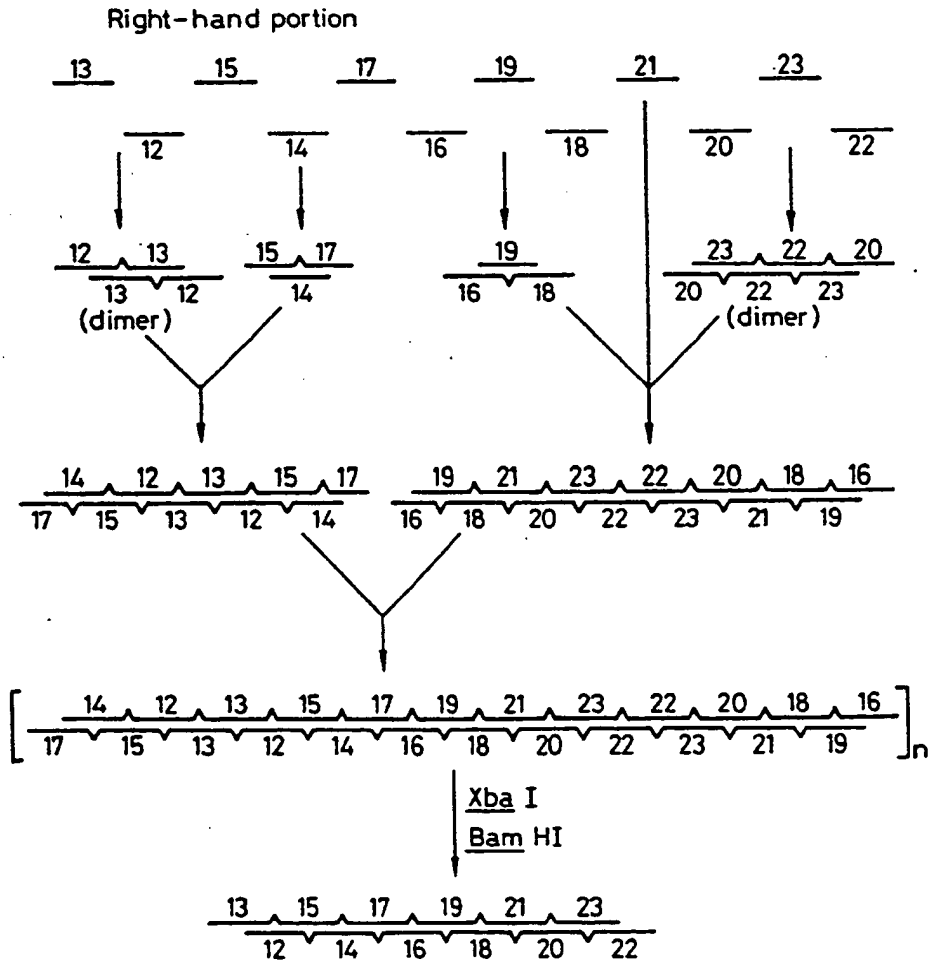


Fig.8 (Part 2 of 2)

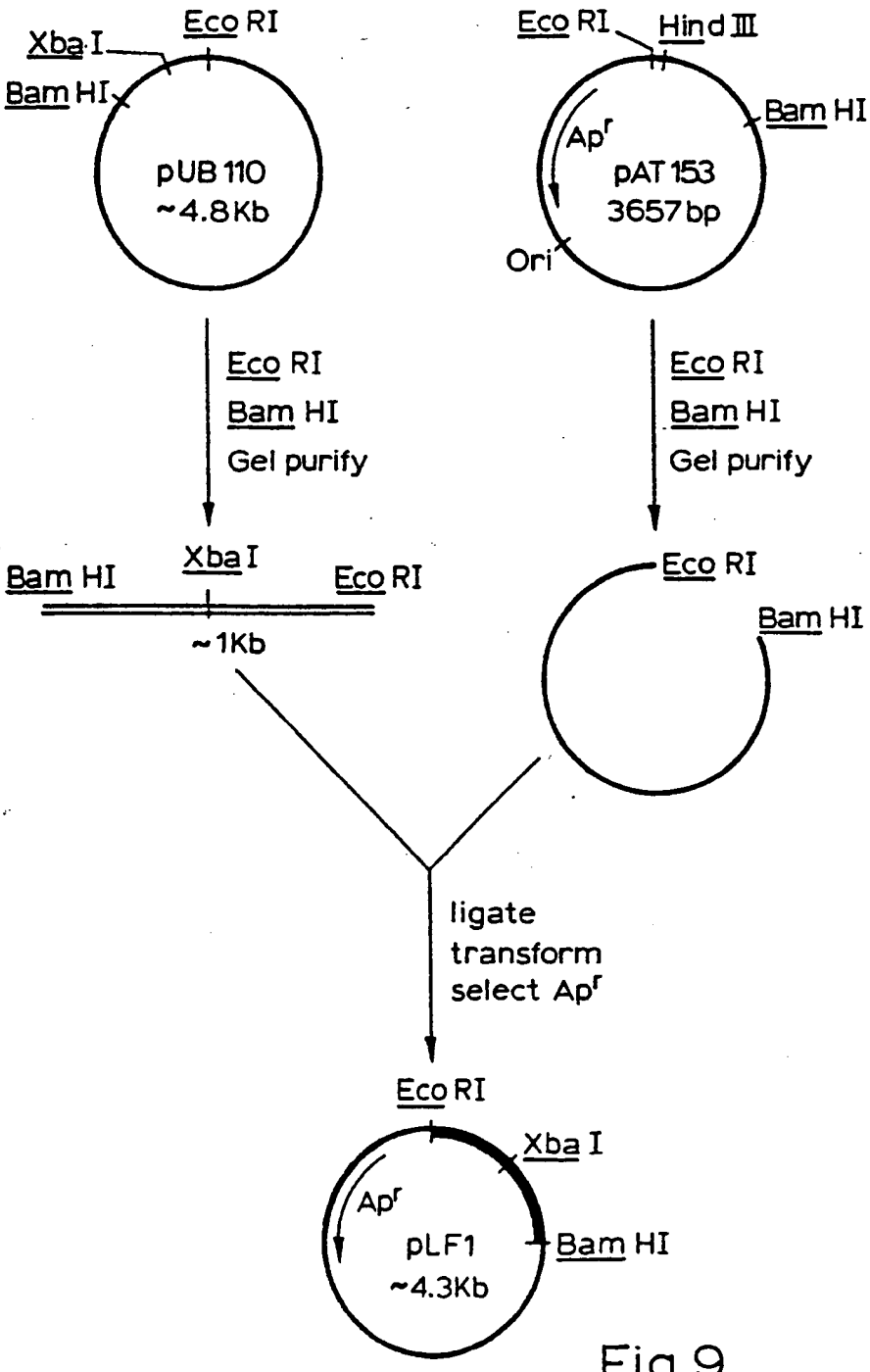

Xba I
Eco RI
Bam HI
pUB 110
~4.8Kb
Eco RI
Hind III
Bam HI
Ap^r
pAT 153
3657bp
Ori
Eco RI
Bam HI
Gel purify
Eco RI
Bam HI
Gel purify
Bam HI
Xba I
Eco RI
~1Kb
Eco RI
Bam HI
ligate
transform
select Ap^r
Eco RI
Xba I
Ap^r
pLF1
~4.3Kb
Bam HI

Fig.9

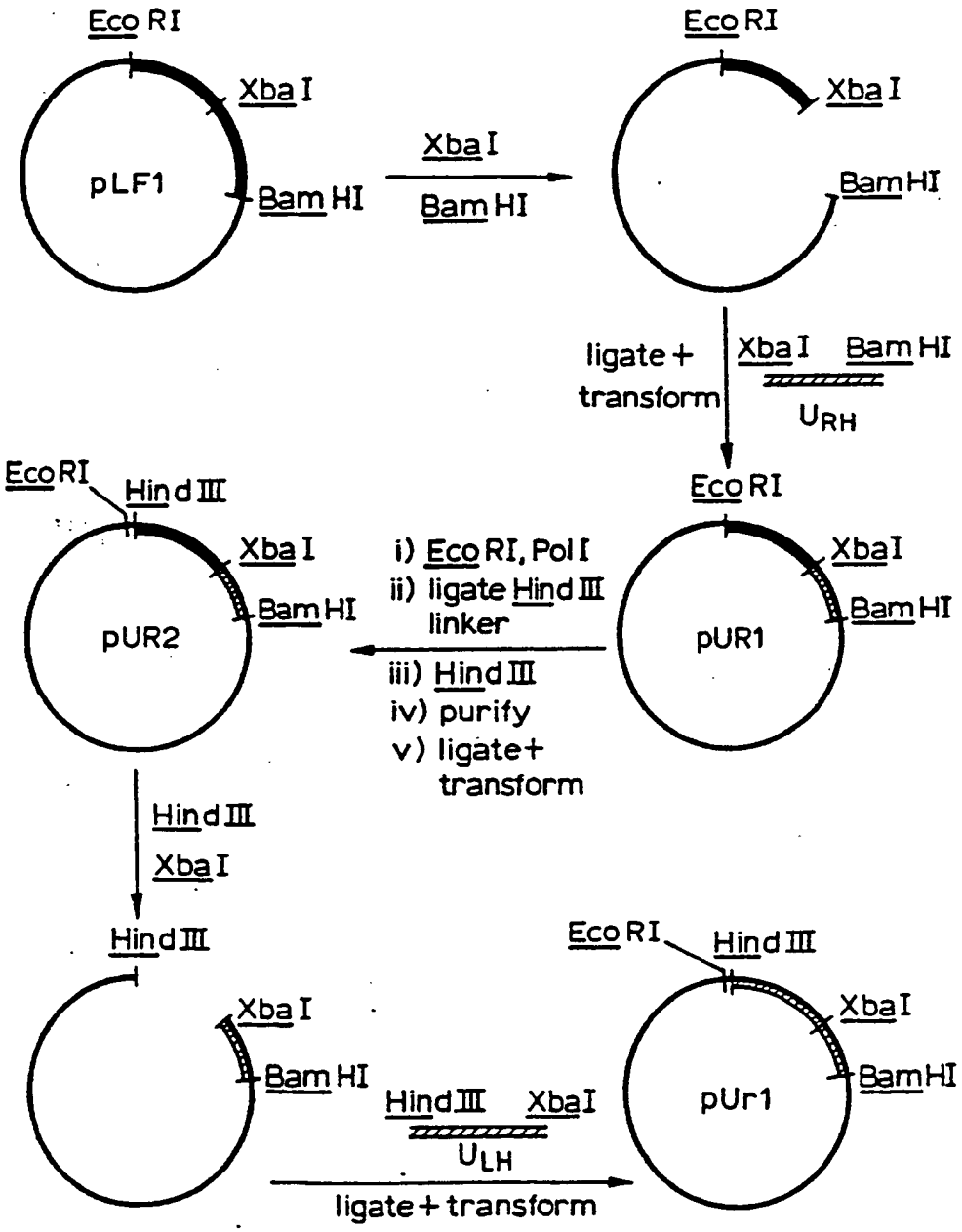

Eco RI
Xba I
pLF1
Bam HI
Xba I
Bam HI
Eco RI
Xba I
Bam HI
ligate +
transform
Xba I
Bam HI
URH
Eco RI
Hind III
Xba I
pUR2
Bam HI
i) Eco RI, Pol I
ii) ligate Hind III
linker
iii) Hind III
iv) purify
v) ligate +
transform
Eco RI
Xba I
pUR1
Bam HI
Hind III
Xba I
Hind III
Xba I
Bam HI
Hind III
Xba I
ULH
ligate + transform
Eco RI
Hind III
Xba I
pUr1
Bam HI

Fig.10

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 046 039**
B2

(12) # NEW EUROPEAN PATENT SPECIFICATION

(45) Date of publication of the new patent specification: **24.01.90**

(51) Int. Cl.[5]: **C 12 N 15/00**, C 12 P 21/02, C 07 H 21/04, C 12 N 1/20

(21) Application number: **81303517.7**

(22) Date of filing: **31.07.81**

(54) **Synthetic urogastrone gene, corresponding plasmid recombinants, transformed cells, production thereof and urogastrone expression.**

(30) Priority: **05.08.80 GB 8025440**

(43) Date of publication of application:
17.02.82 Bulletin 82/07

(45) Publication of the grant of the patent:
14.01.87 Bulletin 87/03

(45) Mention of the opposition decision:
24.01.90 Bulletin 90/04

(84) Designated contracting states:
BE CH DE FR GB IT LI NL SE

(56) References cited:
EP-A- 0 001 931
Dictionary of Microbiology, John Wiley and Sons, 1980, page 244
Fiers et al., Nature, Vol. 260, April 8, 1976, pp. 500-501
Itakura et al., Science, Vol. 196, December 9, 1977, pp. 1056-1063
Crea et al., PNAS USA, Vol. 75, 1978, pp. 5765-5769
Goeddel et al., PNAS USA, Vol. 76, 1979, pp. 106-110
Smith et al., Nuc. Acids Res., Vol. 10, 1982, p. 4467
Grantham et al., Nuc Acids Res., Vol. 9, 1981, p. 43
Gouy et al., Nuc Acids Res., Vol. 10. 1982, p. 7055
Ikemura, J. Mol. Biol, Vol. 146, 1981, pp. 1-21
Storur et al., Nuc. Acids Res., Vol. 10, 1982, p. 2971
Cell, Vol. 27, 1981, p. 533
Cell, Vol. 29, 1982, p. 727
PNASS, Vol. 79, 1982, p. 238
Gheysen et al. Gene, Vol. 17, 1982. pp. 55-63
Iserentant et al., Gene, Vol. 9, 1980, pp. 1-12
Buckhingham and Carrier, Interaction of translational and transcriptional controls in the regulation of gene expression, 1982, 215, Ed. Grumber-Manago et al., Elsevier Science Publ.
Gold et al., Ann, Rev. Microbiol., Vol., 35, 1981, pp. 365-403
Watson, Molecular Biology of the Gene, 3rd Ed., pp. 443-445
PNAS, Vol. 80, 1983, p. 653

(73) Proprietor: **G.D. Searle & Co.**
P.O. Box 1045
Skokie Illinois 60076 (US)

(72) Inventor: **Eaton, Michael Anthony William**
43, Perry Lane Bledlow
Aylesbury Buckinghamshire (GB)
Inventor: **Smith, John Craig**
2 Edgecote House Amersham Hill
High Wycombe Buckinghamshire (GB)
Inventor: **Doel, Michael Terence**
Woodbury Wycombe Road
Studley Green Buckinghamshire (GB)
Inventor: **Lilley, David Malcolm James**
"The Haven" 15 commonside Downley
High Wycombe Buckinghamshire (GB)
Inventor: **Carey, Norman Henry**
Russells Close High Street
Chinnor Oxfordshire (GB)
Inventor: **Bell, Leslie David**
13 Brookside
Thame Oxfordshire (GB)

(74) Representative: **Smith, Sydney et al**
Elkington and Fife Beacon House 113 Kingsway
London WC2B 6PP (GB)

EP 0 046 039 B2

Rosenberg and Court, Ann. Rev. Genet. Vol. 13, 1979, p. 319
Gene Cloning in organisms other than E.Coli, 19-29, Ed. Hofschneider, P.H., and Goebel W., Springer Verlag
Plasmid, Vol. 6, 1981, p. 193

Description

This invention relates to a synthetic urogastrone gene, to corresponding plasmid recombinants and transformed cells, to the production thereof and to urogastrone expression.

Urogastrone is a polypeptide hormone (protein) synthesised in the duodenum and in the salivary glands of normal humans, (see, for example, Heitz et al. (1978), GUT, 19, 408-413). Urogastrone suppresses the secretion of gastric acid and promotes cell growth (see, for example, Elder et al. (1975), Gut, 16, 887-893). Therefore, it has an application in the treatment of ulcers and in the promotion of wound healing. Urogastrone is excreted in small amounts in human urine and may be isolated therefrom. There exists, however, a need for a more viable commercial production thereof and such is provided according to the present invention.

Urogastrone is known to consist of 53 amino acids in the following sequence :

Asn Ser Asp Ser Glu Cys Pro Leu Ser His Asp-
Gly Tyr Cys Leu His Asp Gly Val Cys Met Tyr-
Ile Glu Ala Leu Asp Lys Tyr Ala Cys Asn Cys-
Val Val Gly Tyr Ile Gly Glu Arg Cys Gln Tyr-
Arg Asp Leu Lys Trp Trp Glu Leu Arg

(see, for example, Gregory H. and Preston B. M. (1977), Int. J. Peptide Protein Res., 9, 107-118).

From the above amino acid sequence, a corresponding synthetic gene sequence has been invented, subject to a number of specific non-obvious criteria, and oligonucleotide blocks synthesised which, when assembled, form a synthetic gene coding for urogastrone. The blocks have been hybridised and ligated in pre-determined stages to construct the urogastrone gene in two portions. These have been cloned in two operations into a new specifically-designed chimeric E. coli/S. aureus vector so as to produce a full length urogastrone gene flanked only by E. coli plasmid DNA. The gene has been excised from this recombinant and re-cloned into vectors specifically designed to maximise expression of the gene in E. coli, under the control of the promoter obtained from the E. coli tryptophan operon. A protein resembling human urogastrone has thus been expressed in E. coli.

From the above amino acid sequence, because of the degeneracy of the genetic code, it is possible to predict numerous nucleotide sequences which would code for the protein.

In the inventive determination of an optimum sequence from the large number of possibilities, several non-obvious criteria have been observed. Firstly, trinucleotide codons should be used which are acceptable or preferable in the cells to be used, in particular E. coli. Secondly, it was decided that it was desirable to have different restriction enzyme recognition sites at the termini of the molecule so as to allow insertion into a plasmid in a desired orientation. Moreover, it was decided to select sites which allowed the use of well-understood cloning vectors, such as pBR322 (see, for example, Bolivar F. et al. (1977), Gene, 2, 95-113). In fact, Hind III and Bam HI sites were selected and introduced at the 5' and 3' ends, respectively. Thirdly, it was thought desirable to introduce a series of restriction endonuclease recognition sites strategically placed along the molecule to enable the gene to be specifically disected to aid characterisation and, possibly mutagenesis. Also, this measure allowed the two portions of the molecule to be cloned in stages. In particular, an Xba I site was introduced at a central location in the gene. Fourthly, the synthesis should not be unnecessarily complicated and illegitimate cross-hybridisations should be minimised in order to facilitate gene assembly. Using a computer, energies of interactions arising from all possible approximations were calculated (see, for example, Tinoco Jr. I et al. (1971), Nature, 230, 362-367 ; Powers G. J. et al. (1975), JACS, 97, 875-899, so that stable off-diagonal interactions might be avoided where possible. Fifthly, since the protein ultimately expressed in bacterial cells will be in the form of a fusion product, it was desirable to have a means of cleaving the urogastrone portion from such a fusion product. Since urogastrone is known to be insensitive to trypsin (see, for example, Gregory and Preston, loc. cit.), the codons specifying the dipeptide lys-lys were introduced near the end of the gene corresponding to the urogastron N-terminus in order to serve as a substrate for tryptic proteolysis.

The present invention relates to a synthetic gene characterised in that it codes for the expression of urogastrone or an equivalent thereof or a sub-unit thereof in a bacterial cell and comprises the following sequence or sub-unit thereof :

5' A A T T C C G A T A G C G A G T G T C C T C T G-
3' T T A A G G C T A T C G C T C A C A G G A G A C-
A G T C A C G A T G G T T A C T G T C T A C A T G A C G G-
T C A G T G C T A C C A A T G A C A G A T G T A C T G C C-
C G T C T G T A T G T A T A T T G A G G C T C T A G A C A-
G C A G A C A T A C A T A T A A C T C C G A G A T C T G T-
A G T A C G C G T G T A A T T G C G T T G T T G G C T A C-
T C A T G C G C A C A T T A A C G C A A C A A C C G A T G-

```
A T C G G T G A G C G C T G T C A G T A T C G A G A T C T-
T A G C C A C T C G C G A C A G T C A T A G C T C T A G A-
G A A A T G G T G G G A A C T T A G A
C T T T A C C A C C C T T G A A T C T
```

In fact, bearing in mind, inter alia, the above-mentioned considerations, for convenience, a slightly longer sequence was selected which is as follows :

```
Coding 5' A G C T T A A A A A G A A T T C C G A T A G C G A G T-
Non-Coding 3' A T T T T T C T T A A G G C T A T C G C T C A-
G T C C T C T G A G T C A C G A T G G T T A C T G T C T-
C A G G A G A C T C A G T G C T A C C A A T G A C A G A-
A C A T G A C G G C G T C T G T A T G T A T A T T G A G-
T G T A C T G C C G C A G A C A T A C A T A T A A C T C-
G C T C T A G A C A A G T A C G C G T G T A A T T G C G-
C G A G A T C T G T T C A T G C G C A C A T T A A C G C-
T T G T T G G C T A C A T C G G T G A G C G C T G T C A-
A A C A A C C G A T G T A G C C A C T C G C G A C A G T-
G T A T C G A G A T C T G A A A T G G T G G G A A C T T-
C A T A G C T C T A G A C T T T A C C A C C C T T G A A-
A G A T A A G
T C T A T T C C T A G
```

The whole of this expanded sequence is shown in Figure 1 of the accompanying drawings and the restriction enzyme recognition sites are shown in Figure 2 of the accompanying drawings.

The present invention also relates to a process for the production of the above said synthetic genes or sub-units thereof characterised in that it comprises the assembly and ligation of a number of oligonucleotide blocks.

Such « blocks » should not be regarded as sub-units of the gene. In the present context, « sub-unit » refers to a sequence which is less than the whole, but which exhibits the desired properties.

It was in fact decided to synthesise a molecule having the above expanded sequence by making 23 synthetic oligonucleotide blocks as illustrated in Figure 3 of the accompanying drawings, which will assemble by single-strand overlaps to give the complete double-stranded nucleotide sequence.

In order to minimise strong illegitimate interactions near the centre of the molecule during assembly, certain blocks of modified sequence have also been synthesised, as illustrated in Figure 4 of the accompanying drawings.

As mentioned above, the synthetic blocks selected are shown in Figure 3 of the accompanying drawings. The blocks may be constructed using known synthesis techniques (see, for example, Agarwal, et al. (1970), Nature, 227, 27-34 ; and Crea et al. (1978), Proc. Natl. Acad. Sci. U.S.A., 75, 5765-5769).

The synthetic methods will now be illustrated with reference to the synthesis of the tetradecanucleotide ApGpTpTpCpCpCpApCpCpApTpTpT.

The methods of building up oligonucleotides from smaller units by successive coupling reactions are well known (see, for example, Hsiung et al. (1979), Nucleic Acid Research, 6, 1371-1385). The completely protected tetradecanucleotide was built up as shown in Figure 5 of the accompanying drawings (wherein, for convenience, protecting groups are not shown).

The condensation reactions indicated by arrows in Figure 5 were carried out by the following procedure exemplified in the synthesis of :

$$\text{DMTr } A^{Bz}p\ G'^{Bu}p\text{-CNEt}$$
$$\vdots \quad \vdots$$
$$\phi Cl \quad \phi Cl$$

as shown in Figure 6 of the accompanying drawings. From 1.1 to 1.5 mmole of the 3'-phosphodiester component (I) was condensed with 1.0 mmole of the 5'-hydroxyl component (II) in anhydrous pyridine in the presence of from 3 to 4.5 mmoles of 2,4,6-triisopropyl-benzene-sulphonyl-tetrazolide. The reaction was left for 1 hour at room temperature or until chromatography on silica TLC plates eluted with 10 % (v/v) methanol in chloroform showed that the hydroxyl component was exhausted. The reaction was quenched with 5 % (w/v) sodium bicarbonate solution and extracted with chloroform. The chloroform extract was dried and loaded onto a reverse phase chromatography column (ODS bonded to 15-25 micron silica). The fully protected dinucleotide product (III) was eluted with a solvent gradient from chloroform : methanol :

water (2 : 6 : 3 v/v) to chloroform : methanol : water (2 : 6 : 0.5 v/v). The product (III) was extracted into chloroform and dried. The final isolated yield was 81 %.

To proceed to further condensations, the terminal protecting group (DMTr or CNEt) was removed selectively using triethylamine in pyridine (CNEt) or a 2 % (w/v) solution of benzene sulphonic acid in chloroform : methanol (DMTr) as shown in Figure 7 of the accompanying drawings.

At the completion of the synthesis, all of the protecting groups were removed by sequential treatment with 0.1 M tetraethylammonium fluoride in THF/pyridine/water (8 : 1 : 1 v/v), ammonia and 80 % acetic acid. The deprotected oligonucleotides were purified by ion-exchange HPLC and sequence analysis was carried out by the method of Wu et al. (1976), Anal. Biochem., 74, 73-93.

The oligomeric blocks of nucleotides were hybridised and ligated (see, for example, Agarwal et al., loc. cit.) in a series of steps, in order to minimise the possibilities for undesirable interactions, leading to the formation of the two portions as shown in Figure 8 of the accompanying drawings. The order of the additions in the assembly scheme was optimised for minimal incorrect ligations and in the case of especially difficult oligomeric blocks, notably 7 and 8, sub-molar quantities were used in order to remove all monomeric units before further additions were made.

In more detail :

Left-hand portion : Blocks 1 and 2 were ligated to form a dimer about the Hind III site. Blocks 4, 5 and 6 and 8, 9 and 10 were also ligated in the first round of the assembly scheme. Molar equivalents were used for all but block 8 where 0.75 molar equivalents were employed.

Block 3 was ligated with 1 + 2, block 7 (0.75 molar equivalent) with 4 to 6 and block 11 with 8 to 10. The 8 + 9 + 10 + 11 assembly has one flush end, hence some blunt-end dimerisation was observed.

1 to 3 and 4 to 7 were ligated and finally 8 to 11 were ligated to the resulting 1 to 7 species.

The dimeric 1 to 11 left-hand portion was then cleaved by Hind III (EC 3.1.23.21) and Xba I (EC 3.1.23.4) to generate the monomeric left-hand portion, with the correct cohesive termini to allow construction of recombinant plasmids.

Right-hand portion : Blocks 12 and 13 were ligated to form a dimer about the Xba I site and blocks 20, 23 and 22 similarly ligated to form a dimer about the Bam HI site. Blocks 14, 15 and 17, and 16, 18 and 19 were also ligated at this stage.

The 12, 13 dimer and the 14, 15, 17 assembly were ligated, as were the 16, 18, 19 assembly with the 20, 22, 23 dimer, where block 21 was used as a joining section.

These two species were then ligated to give an oligomeric molecule which was cleaved by Xba I and Bam HI (EC 3.1.23.6) to give the monomer 12 to 23 species having the correct cohesive termini to allow construction of recombinant plasmids.

The present invention further relates to a plasmid recombinant characterised in that it comprises a plasmid vector having inserted therein at an appropriate insertion site a synthetic gene or a sub-unit thereof as defined above, the plasmid recombinant enabling translation in the correct phase for the mRNA corresponding to the inserted gene or sub-unit thereof and having a bacterial promoter upstream of and adjacent to the insertion site such that the inserted gene or sub-unit thereof is under bacterial promoter control.

A particular vector, designated pLFI, has been specifically designed to afford advantages for the purposes of the present invention. The inventive plasmid pLFI, is a 5K bp plasmid which may be propagated in E. coli and which may be constructed from pBR322 and pUB110 (see, for example, Gryczan, T. J. et al. (1978), J. Bacterial., 134, 318) by inserting the DNA sequence of the S. aureus plasmid pUB110 between the EcoRI and Bam HI sites thereof (comprising approximately 870 bp) between the EcoRI and Bam HI sites of pBR322, thereby replacing that region of pBR322.

The present invention further relates to a process for the production of such a plasmid recombinant characterised in that it comprises inserting such a synthetic gene or a sub-unit thereof as defined above into an appropriate insertion site of an appropriate plasmid vector.

The following illustrates the present invention :

Phosphorylation of oligonucleotide blocks : In each case, 6 μg of oligomer and 60 μCi of [γ-^{32}P] ATP (> 5 000 Ci/mMol) were dried and redissolved in a final buffer concentration of 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 0.25 mM ATP, 10 mM β-mercaptoethanol and incubated with 4 units (1 unit is the amount that catalyses the production of 1 n mole of acid-insoluble ^{32}P after incubation for 30 minutes at 37 °C according to Richardson, C.C., (1972), Progress in Nucleic Acids Research, 2, 815) of T4 polynucleotide kinase (EC 2.7.1.78, Bethesda Research Labs) at 37 °C for 15 minutes. The enzyme was subsequently inactivated by a 5 minute incubation at 100 °C.

Ligation of oligonucleotide blocks : Except where indicated above, equimolar quantities (6 to 18 μg) of oligonucleotide blocks were incubated in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM dithiothreitol (DTT) with 6 units (1 unit is the amount that catalyses the conversion of 1 n mole of ^{32}PPi into (α/β ^{32}P)-ATP in 20 minutes at 37 °C according to Weiss B. et al. (1968), J. Biol. Chem., 243 4543) of T4 DNA ligase (EC 6.5.1.1, Bethesda Research Labs) at 25 °C for from 3 to 16 hours. Ligated DNA was precipitated by addition of 2.5 vol absolute ethanol, collected by centrifugation and redissolved in water.

Purification of ligated species : Ligated oligonucleotide blocks were electrophoresed on 20 % (w/v) polyacrylamide in 90 mM Tris-HCl, pH 8.3, 90 mM boric acid, 2.5 mM EDTA (TBE buffer), and the fragments located by autoradiography. Slices of gel containing the fragments were excised and the DNA

electroeluted at 1 mA in TBE buffer onto 0.5 ml of DEAE cellulose (DE52, Whatman) for a few hours. After extensive washing by 0.1MNH_4OAc, 2 mM $Mg(OAc)_2$, 0.02 % (w/v) SDS, 0.02 mM EDTA (AGEB buffer), the DNA was eluted from the DEAE cellulose by 2 ml of 1.1 M NaCl in AGEB buffer and precipitated by addition of 2.5 vol absolute ethanol.

Construction of chimeric cloning vector pLF1 : In order to facilitate the two-stage cloning of urogastrone, it was desirable to construct a vector or vectors having Hind III, Xba I and Bam HI cleavage sites, see Figure 9 of the accompanying drawings. However, no readily available E. coli plasmids possess Xba I sites. It was surprisingly noticed that the S. aureus plasmid pUB110, which cannot be propagated in E. Coli (see, for example, Keggins K. M. et al. (1978), Proc. Natl. Acad. Sci. U.S.A., 75, 1423), contains a region of about 1kbp of DNA bounded by sites for EcoRI and Bam HI having an approximately central Xba I site. Therefore, pUB110 was cleaved with EcoRI (EC 3.1.4.32) and Bam Hi and the DNA fragments electrophoresed on 5 % (w/v) polyacrylamide. The approximately 1 kbp EcoRI/Bam HI fragment was removed by electroelution from the excised gel slice onto DEAE cellulose, eluted by 1.1 M NaCl and ethanol precipitated. The E. coli plasmid pAT153 (see, for example, Twigg, A. J., and Sherratt D. (1980), Nature, 283, 216-218) was also cleaved with EcoRI and Bam HI, and the 3282 bp fragment purified by electroelution as above. These two purified DNA fragments were ligated in equimolar quantities using 10 units of T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT at 15° for 18 hours. After ethanol precipitation, the ligated DNA was dissolved in 100 µl 10 mM Tris pH 7.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$ and transformed into E. coli K12 HB101 (genotype gal^-, lac^-, ara^-, pro^-, arg^-, str^r, rec A^-, r_k^-, M_k^- ; see, for example, Boyer H. W ; and Roullard-Dussoix D., J. Mol. Biol., 41, 459-472) using known methods (see, for example, Cohen et al. (1972), Proc. Natl. Acad. Sci. U.S.A., 69, 2110-2114) and transformants resistant to 100 µg/ml ampicillin selected. Several transformants were analysed further by restriction enzyme cleavage of plasmid DNA, and one full length clone, designated pLF1, selected for further use as a cloning vector. At a later stage, an additional Hind III restriction sequence was created at the EcoRI site.

Cloning of the synthetic urogastrone gene in pLF1 : The two portions of the assembled urogastrone gene were cloned in two transformation stages, as illustrated in Figure 10 of the accompanying drawings.

Right-hand portion : The longer fragment of Xba I, Bam HI-cleaved pLF1 was purified by electroelution as above and ligated to a large excess of the right-hand portion assembly of the urogastrone gene using 10 units T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT at 15 °C for 18 hours. This was transformed into E. coli as before, with selection for 100 µg/ml ampicillin. Several transformants were selected for plasmid analysis by restriction enzyme cleavage analysis and one clone, designated pUR1, used for further cloning. The sequence of the inserted urogastrone gene DNA was confirmed by chemical degradation analysis (see, for example, Maxam, A., and Gilbert, W., (1977), Proc. Natl. Acad. Sci. U.S.A., 74, 560-4).

In order to have a Hind III site for ligation to the 5' end of the left-hand portion of the gene, the EcoRI site was modified as follows : pUR1 was cleaved with EcoRI and the resulting recessed ends filled using 5 units (1 unit is the amount that incorporates 10 n moles of total nucleotides into an acid-precipitable fraction in 30 minutes at 37 °C using poly-d(A-T) as primer according to Richardson, C.C. et al. (1964), J. Biol. Chem., 239, 222) of DNA polymerase I (EC 2.7.7.7) in 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM β-mercaptoethanol with 0.25 mM of each deoxynucleotide for 30 minutes at 15 °C.

To the resulting flush ends, were ligated a large excess of synthetic Hind III linkers (Collaborative Research) by the known blunt end ligation procedure (see, for example, Ullrich A. et al. (1977), Science 196, 1313-1319). The DNA was ligated using 6 units of T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT at 15 °C for 18 hours.

After Hind III restriction cleavage and purification, the full length plasmid was religated and transformed into E. coli as above, with selection for 100 µg/ml ampicillin. A transformant having a Hind III site, designated pUR2, was selected for further cloning.

Left-hand portion : pUR2 was cleaved with Hind III and XbaI restriction enzymes and the longer fragment purified by electroelution as above. This was ligated to a two-fold molar excess of the assembled left-hand portion of the urogastrone gene using 6 units T4 DNA ligase in 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM ATP, 20 mM DTT for 18 hours at 15 °C. This was transformed into E. coli K12 MRC 8 (genotype dap 103 hsd R met BI glm 533 upp 1 dap 101 sup E thy A 103 deo rec A1) with selection for 100 µg/ml ampicillin. Several transformants were selected for plasmid analysis by restriction enzyme cleavage. One clone, designated pUR1 was used for further characterisation and cloning. The sequence of the full urogastrone gene was confirmed by chemical degradation analysis. It should be noted that pUR1 has no remaining S aureus DNA sequence present.

The present invention also relates to a bacterial cell, in particular an E. coli cell, characterised in that it comprises inserted therein a plasmid recombinant as defined above.

The present invention further relates to a process for the production of such a bacterial cell characterised in that it comprises inserting a plasmid recombinant as defined above into a bacterial cell.

Expression of urogastrone in E. coli : The urogastrone gene insert was cleaved from pUR1 by Hind III and Bam HI cleavage and purified by polyacrylamide gel electrophoresis and electroelution as above. This fragment was ligated to Hind III, Bam HI-cleaved pWT121 and pWT221, (see, for example, Tacon, W.C.A. et al. (1980), Molec. Gen. Genet. 177, 427) and the recombinant molecules used to transform E. coli MRC 8,

(see for example, Emtage, J. S. et al. (1980), Nature, 283, 171-174), with selection for 100 μg/ml ampicilin. Transformants containing full length urogastrone genes were characterised by restriction enzyme cleavage analysis and DNA was purified by isopycnic centrifugation in caesium chloride.

Expression of urogastrone-like fusion protein was induced by growth of cells in L-broth (luria broth : 1 % (w/v) bacto tryptone, 0.5 % (w/v) bacto yeast extract, 0.5 % (w/v) NaCl, 0.2 % (w/v) glucose, 0.004 % (w/v) thymine, pH 7) containing 100 μg/ml ampicillin to an A600 nm of 0.3. Following centrifugation, the cells were washed and resuspended in M9 medium lacking tryptophan, but containing 20 μg/ml 3 β-indole acrylic acid. The cells were incubated at 37 °C for 4 hours. Under these conditions maximal tryptophan promoter activity is known to occur (see Tacon et al. loc cit), and hence expression of the urogastrone fusion protein.

The present invention also relates to a process for the production of urogastrone or an equivalent thereof or a sub-unit thereof characterised in that it comprises culturing a bacterial cell as defined above and recovering expressed protein.

The present invention further relates to a fused polypeptide characterised in that it comprises urogastrone or an equivalent thereof or a sub-unit thereof covalently bonded with all or part of a gene of the tryptophan operon, e. g. trp E, and in that it is obtained by culturing a bacterial cell as defined above and recovering expressed protein.

As mentioned above, urogastrone has an application in the treatment of ulcers, and also in other instances where the growth promoting activity thereof would be beneficial, for example, in wound healing. Conventional administration forms may be used, the active material being used in an effective amount, for example from 0.1 to 1.0 μg/kg body weight, preferably from 0.125 to 0.5 μg/kg body weight, more preferably about 0.25 μg/kg body weight, optimally together with a conventional pharmaceutically-acceptable carrier, diluent or adjuvant.

Claims

1. A synthetic gene characterised in that it codes for the expression of urogastrone or an equivalent thereof or a sub-unit thereof in a bacterial cell and comprises the following sequence or a sub-unit thereof :

```
Coding: 5' A A T T C C G A T A G C G A G T G T C C T C T-
Non-coding: 3'  T T A A G G C T A T C G C T C A C A G G A G A-
G A G T C A C G A T G G T T A C T G T C T A C A T G A C G G-
C T C A G T G C T A C C A A T G A C A G A T G T A C T G C C-
C G T C T G T A T G T A T A T T G A G G C T C T A G A C A A-
G C A G A C A T A C A T A T A A C T C C G A G A T C T G T T-
G T A C G C G T G T A A T T G C G T T G T T G G C T A C A T-
C A T G C G C A C A T T A A C G C A A C A A C C G A T G T A-
C G G T G A G C G C T G T C A G T A T C G A G A T C T G A A-
G C C A C T C G C G A C A G T C A T A G C T C T A G A C T T-
A T G G T G G G A A C T T A G A
T A C C A C C C T T G A A T C T
```

2. A synthetic gene as claimed in Claim 1 characterised in that it comprises the following sequence or a sub-unit thereof :

```
Coding: 5' A G C T T A A A A A G A A T T C C G A T A G-
Non-coding: 3' A T T T T T C T T A A G G C T A T C-
C G A G T G T C C T C T G A G T C A C G A T G G T T A C T-
G C T C A C A G G A G A C T C A G T G C T A C C A A T G A-
G T C T A C A T G A C G G C G T C T G T A T G T A T A T T-
C A G A T G T A C T G C C G C A G A C A T A C A T A T A A-
G A G G C T C T A G A C A A G T A C G C G T G T A A T T G-
C T C C G A G A T C T G T T C A T G C G C A C A T T A A C-
C G T T G T T G G C T A C A T C G G T G A G C G C T G T C-
G C A A C A A C C G A T G T A G C C A C T C G C G A C A G-
A G T A T C G A G A T C T G A A A T G G T G G G A A C T T-
T C A T A G C T C T A G A C T T T A C C A C C C T T G A A-
A G A T A A G
T C T A T T C C T A G
```

3. A process for the production of a synthetic gene or a sub-unit thereof as claimed in Claim 1 or Claim 2 characterised in that it comprises the assembly and ligation of a number of oligoncleotide blocks.

4. A plasmid recombinant characterised in that it comprises a plasmid vector having inserted therein

at an appropriate insertion site a synthetic gene or a sub-unit thereof as claimed in Claim 1 or Claim 2, the plasmid recombinant enabling translation in the correct phase for the mRNA corresponding to the inserted gene or sub-unit thereof and having a bacterial promoter upstream of and adjacent to the insertion site such that the inserted gene or sub-unit thereof is under bacterial promoter control.

5. A plasmid recombinant as claimed in Claim 4 characterised in that it comprises as the plasmid vector pLFl.

6. A process for the production of a plasmid recombinant as claimed in Claim 4 characterised in that it comprises inserting a synthetic gene or a sub-unit thereof as claimed in Claim 1 or Claim 2 into an appropriate insertion site of an appropriate plasmid vector.

7. A bacterial cell characterised in that it comprises inserted therein a plasmid recombinant as claimed in Claim 4 or Claim 5.

8. A process for the production of a bacterial cell as claimed in Claim 7 characterised in that it comprises inserting a plasmid recombinant as claimed in Claim 4 or Claim 5 into a bacterial cell.

9. A process for the production of urogastrone or an equivalent thereof or a sub-unit thereof characterised in that it comprises culturing a bacterial cell as claimed in Claim 7 and recovering expressed protein.

10. A fused polypeptide characterised in that it comprises urogastrone or an equivalent thereof or a sub-unit thereof covalently bonded with all or part of a gene of the tryptophan operon and in that it is obtained by culturing a bacterial cell as claimed in Claim 7 and recovering expressed protein.

11. A fused polypeptide as claimed in Claim 10 wherein the gene of the tryptophan operon is the trp E gene.

Patentansprüche

1. Synthetisches Gen, dadurch gekennzeichnet, daß es die Expression für Urogastron oder ein Äquivalent desselben oder eine Untereinheit desselben in einer Bakterienzelle codiert und die folgende Sequenz oder eine Untereinheit derselben umfaßt :

```
codierend : 5' A A T T C C G A T A G C G A G T G T C C T C T-
nicht codierend : 3'  T T A A G G C T A T C G C T C A C A G G A G A-
   G A G T C A C G A T G G T T A C T G T C T A C A T G A C G G-
   C T C A G T G C T A C C A A T G A C A G A T G T A C T G C C-
   C G T C T G T A T G T A T A T T G A G G C T C T A G A C A A-
   G C A G A C A T A C A T A T A A C T C C G A G A T C T G T T-
   G T A C G C G T G T A A T T G C G T T G T T G G C T A C A T-
   C A T G C G C A C A T T A A C G C A A C A A C C G A T G T A-
   C G G T G A G C G C T G T C A G T A T C G A G A T C T G A A-
   G C C A C T C G C G A C A G T C A T A G C T C T A G A C T T-
   A T G G T G G G A A C T T A G A
   T A C C A C C C T T G A A T C T
```

2. Synthetisches Gen nach Anspruch 1, dadurch gekennzeichnet, daß es die folgende Sequenz oder eine Untereinheit derselben umfaßt :

```
codierend : 5' A G C T T A A A A A G A A T T C C G A T A G-
nicht codierend : 3' A T T T T T C T T A A G G C T A T C-
   C G A G T G T C C T C T G A G T C A C G A T G G T T A C T-
   G C T C A C A G G A G A C T C A G T G C T A C C A A T G A-
   G T C T A C A T G A C G G C G T C T G T A T G T A T A T T-
   C A G A T G T A C T G C C G C A G A C A T A C A T A T A A-
   G A G G C T C T A G A C A A G T A C G C G T G T A A T T G-
   C T C C G A G A T C T G T T C A T G C G C A C A T T A A C-
   C G T T G T T G G C T A C A T C G G T G A G C G C T G T C-
   G C A A C A A C C G A T G T A G C C A C T C G C G A C A G-
   A G T A T C G A G A T C T G A A A T G G T G G G A A C T T-
   T C A T A G C T C T A G A C T T T A C C A C C C T T G A A-
   A G A T A A G
   T C T A T T C C T A G
```

3. Verfahren zur Erzeugung eines synthetischen Gens oder einer Untereinheit desselben nach Anspruch 1 oder Anspruch 2, dadurch gekennzeichnet, daß es den Zusammenbau und die Verknüpfung einer Anzahl von Oligonucleotid-Blöcken umfaßt.

4. Rekombinantes Plasmid, dadurch gekennzeichnet, daß es einen Plasmid-Vektor umfaßt, der ein synthetisches Gen oder eine Unterheit desselben nach Anspruch 1 oder Anspruch 2, eingefügt an einer

geeigneten Insertionsstelle, aufweist, wobei das rekombinante Plasmid die Translation in der korrekten Phase für die dem eingefügten Gen oder der eingefügten Untereinheit desselben entsprechende mRNA ermöglicht und einen solchen Bakterien-Promotor von (stromaufwärts von) der Insertionsstelle und dieser benachbart aufweist, daß das eingefügte Gen oder die Untereinheit desselben der Steuerung durch den Bakterien-Promotor unterliegen.

5. Rekombinantes Plasmid nach Anspruch 4, dadurch gekennzeichnet, daß es als den Plasmid-Vektor pLFI enthält.

6. Verfahren zur Erzeugung eines rekombinanten Plasmids nach Anspruch 4, dadurch gekennzeichnet, daß es das Einfügen eines synthetischen Gens oder einer Untereinheit desselben nach Anspruch 1 oder Anspruch 2 an einer geeigneten Insertionsstelle eines geeigneten Plasmid-Vektors umfaßt.

7. Bakterienzelle, dadurch gekennzeichnet, daß sie darin eingefügt ein rekombinantes Plasmid nach Anspruch 4 oder Anspruch 5 enthält.

8. Verfahren zur Erzeugung einer Bakterienzelle nach Anspruch 7, dadurch gekennzeichnet, daß es das Einfügen eines rekombinanten Plasmids nach Anspruch 4 oder Anspruch 5 in eine Bakterienzelle umfaßt.

9. Verfahren zur Erzeugung von Urogastron oder eines Aquivalents desselben oder einer Untereinheit desselben, dadurch gekennzeichnet, daß eine Bakterienzelle nach Anspruch 7 kultiviert wird und das exprimierte Protein isoliert wird.

10. Fusioniertes Polypeptid, dadurch gekennzeichnet, daß es Urogastron oder ein Äquivalent desselben oder eine Untereinheit desselben kovalent gebunden mit dem gesamten Gen oder einem Teil des Gens des Tryptophan-Operons umfaßt und daß es durch Kultivieren einer Bacterienzelle nach Anspruch 7 und Isolierung des exprimierten Proteins erhalten wird.

11. Fusioniertes Polypeptid nach Anspruch 10, dadurch gekennzeichnet, daß das Gen des Tryptophan-Operons das trp-E-Gen ist.

Revendications

1. Gène synthétique caractérisé en ce qu'il code pour l'expression d'urogastrone ou un équivalent de cette dernière ou une sous-unité de cette dernière dans une cellule bactérienne et comprend la séquence suivante ou une sous-unité de cette dernière :

```
Codant: 5' A A T T C C G A T A G C G A G T G T C C T C T-
Non-codant: 3' T T A A G G C T A T C G C T C A C A G G A G A-
G A G T C A C G A T G G T T A C T G T C T A C A T G A C G G-
C T C A G T G C T A C C A A T G A C A G A T G T A C T G C C-
C G T C T G T A T G T A T A T T G A G G C T C T A G A C A A-
G C A G A C A T A C A T A T A A C T C C G A G A T C T G T T-
G T A C G C G T G T A A T T G C G T T G T T G G C T A C A T-
C A T G C G C A C A T T A A C G C A A C A A C C G A T G T A-
C G G T G A G C G C T G T C A G T A T C G A G A T C T G A A-
G C C A C T C G C G A C A G T C A T A G C T C T A G A C T T-
A T G G T G G G A A C T T A G A
T A C C A C C C T T G A A T C T
```

2. Gène synthétique selon la revendication 1, caractérisé en ce qu'il comprend la séquence suivante ou une sous-unité de cette dernière :

```
Codant: 5' A G C T T A A A A A G A A T T C C G A T A G-
Non-codant: 3' A T T T T T C T T A A G G C T A T C-
C G A G T G T C C T C T G A G T C A C G A T G G T T A C T-
G C T C A C A G G A G A C T C A G T G C T A C C A A T G A-
G T C T A C A T G A C G G C G T C T G T A T G T A T A T T-
C A G A T G T A C T G C C G C A G A C A T A C A T A T A A-
G A G G C T C T A G A C A A G T A C G C G T G T A A T T G-
C T C C G A G A T C T G T T C A T G C G C A C A T T A A C-
C G T T G T T G G C T A C A T C G G T G A G C G C T G T C-
G C A A C A A C C G A T G T A G C C A C T C G C G A C A G-
A G T A T C G A G A T C T G A A A T G G T G G G A A C T T-
T C A T A G C T C T A G A C T T T A C C A C C C T T G A A-
A G A T A A G
T C T A T T C C T A G
```

3. Procédé de production d'un gène synthétique ou d'une sous-unité de ce dernier selon la revendication 1 ou la revendication 2, caractérisé en ce qu'il comprend l'assemblage et la liaison d'un

certain nombre de fragments d'oligonucléotides.

4. Plasmide recombinant caractérisé en ce qu'il comprend un plasmide vecteur dans lequel est inséré en un site d'insertion approprié un gène synthétique ou une sous-unité de ce dernier selon la revendication 1 ou la revendication 2, le plasmide recombinant permettant la traduction en la phase correcte pour l'ARN-m correspondant au gène inséré ou à la sous-unité de ce dernier et ayant un promoteur bactérien en amont et adjacent par rapport au site d'insertion, tel que le gène inséré ou la sous-unité de ce dernier se trouve sous le contrôle du promoteur bactérien.

5. Plasmide recombinant selon la revendication 4, caractérisé en ce qu'il comprend en tant que plasmide vecteur pLFI.

6. Procédé de production d'un plasmide recombinant selon la revendication 4, caractérisé en ce qu'il comprend l'insertion d'un gène synthétique ou d'une sous-unité de ce dernier selon la revendication 1 ou la revendication 2 dans un site d'insertion approprié d'un plasmide vecteur approprié.

7. Cellulose bactérienne caractérisée en ce que s'y trouve inséré un plasmide recombinant selon la revendication 4 ou la revendication 5.

8. Procédé de production d'une cellule bactérienne selon la revendication 7, caractérisé en ce qu'il comprend l'insertion d'un plasmide recombinant selon la revendication 4 ou la revendication 5 dans une cellule bactérienne.

9. Procédé de production d'urogastrone ou d'un équivalent de cette dernière ou d'une sous-unité de cette dernière, caractérisé en ce qu'il comprend la culture d'une cellule bactérienne selon la revendication 7 et la récupération de la protéine exprimée.

10. Polypeptide fusionné caractérisé en ce qu'il comprend l'urogastrone ou un équivalent de cette dernière ou une sous-unité de cette dernière liée à tout ou partie d'un gène de l'opéron tryptophane et en ce qu'il est obtenu en cultivant une cellule bactérienne selon la revendication 7 et en récupérant la protéine exprimée.

11. Polypeptide fusionné selon la revendication 10, caractérisé en ce que le gène de l'opéron tryptophane est le gène trp E.

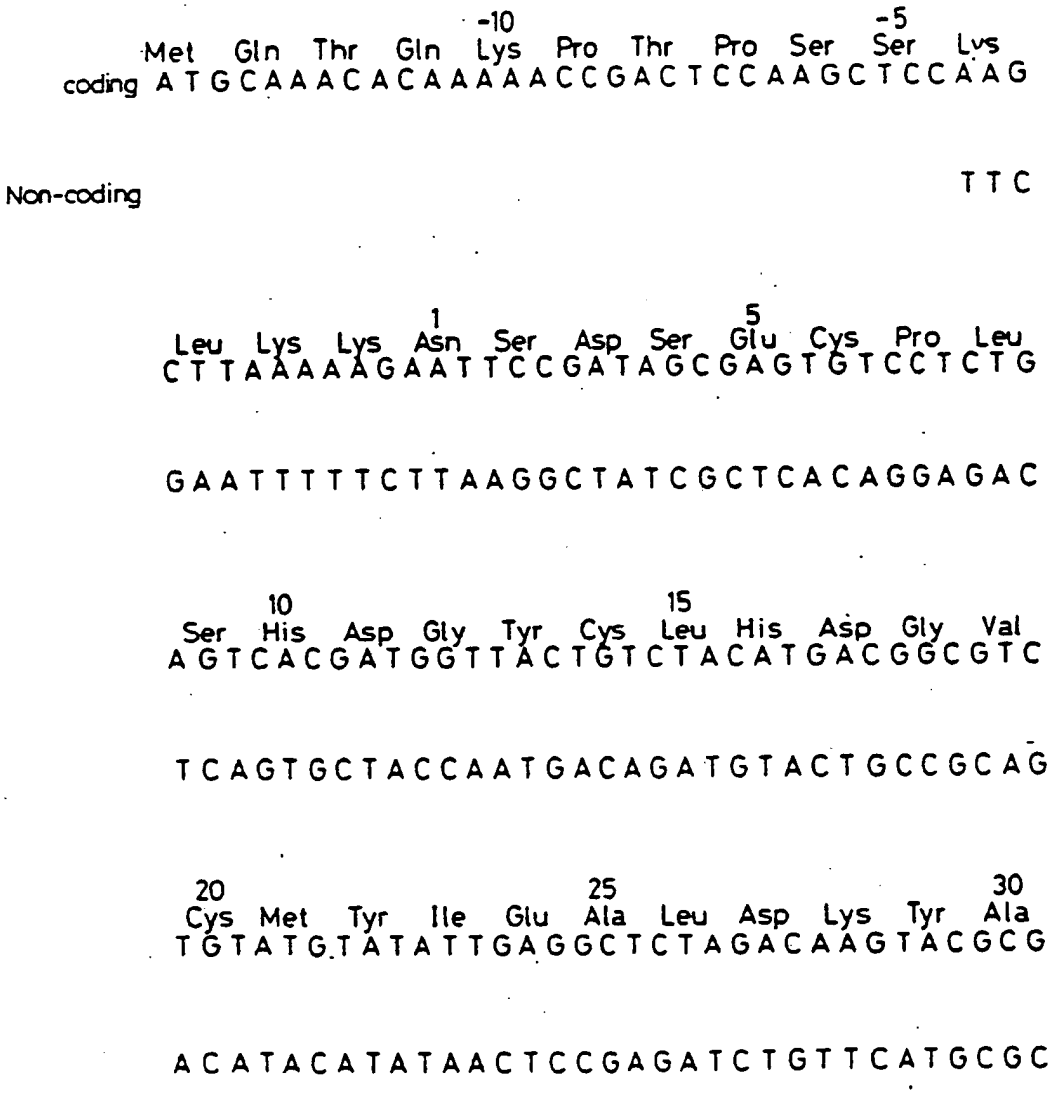


Fig. 1 (Part 1 of 2)

```
                                35                                  40
Cys  Asn  Cys  Val  Val  Gly  Tyr  Ile  Gly  Glu  Arg
TGTAATTGCGTTGTTGGCTACATCGGTGAGCGC

ACATTAACGCAACAACCGATGTAGCCACTCGCG

Cys  Gln  Tyr  Arg  Asp  Leu  Lys  Trp  Trp  Glu  Leu
TGTCAGTATCGAGATCTGAAATGGTGGGAACTT

ACAGTCATAGCTCTAGACTTTACCACCCTTGAA

                 55
Arg  Ter  Gly  Ser
AGATAAGGATCC coding

TCTATTCCTAGG Non-coding
```

Fig.1 (Part 2 of 2)

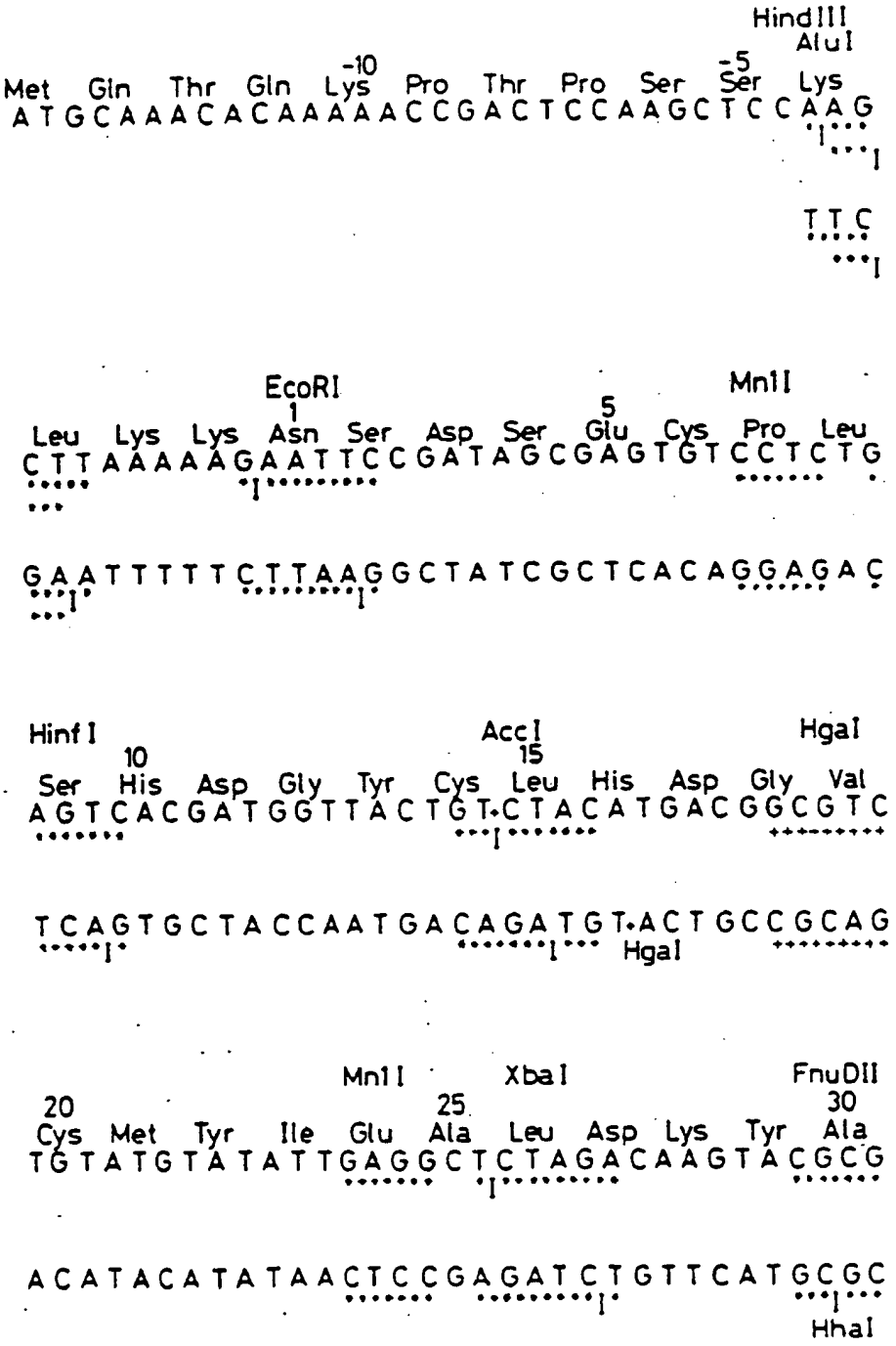


Fig. 2 (Part 1 of 2)

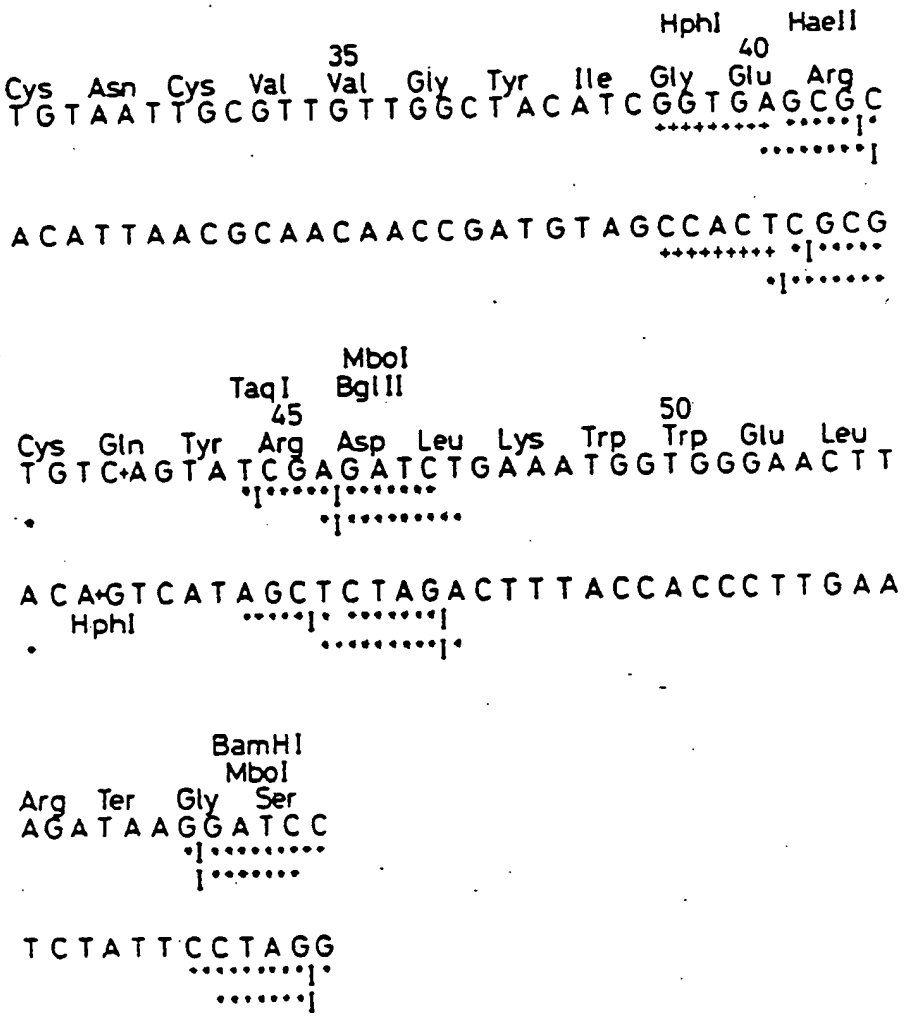


Fig. 2 (Part 2 of 2)

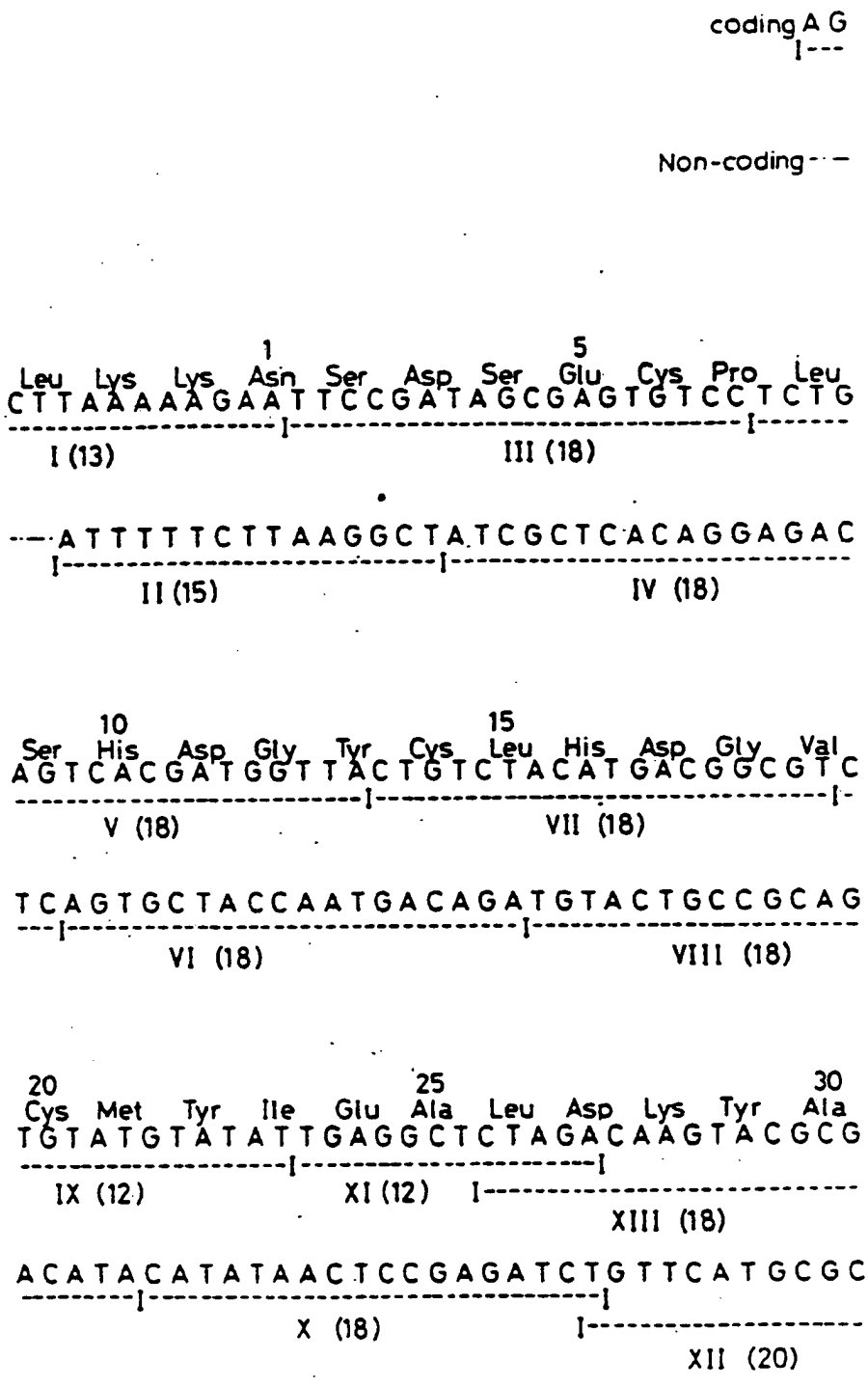


Fig. 3 (Part 1 of 2)

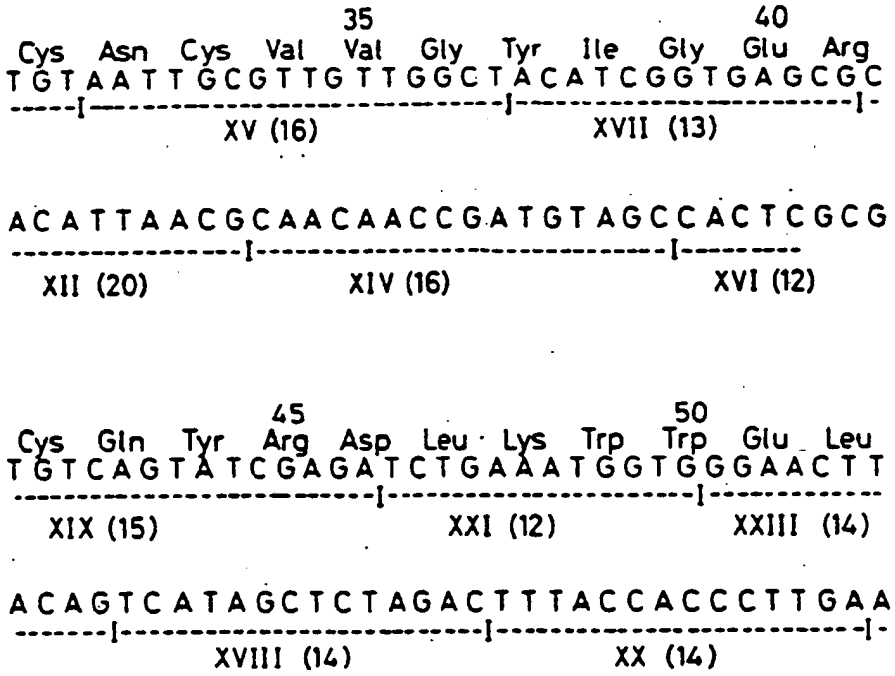


Fig.3 (Part 2 of 2)

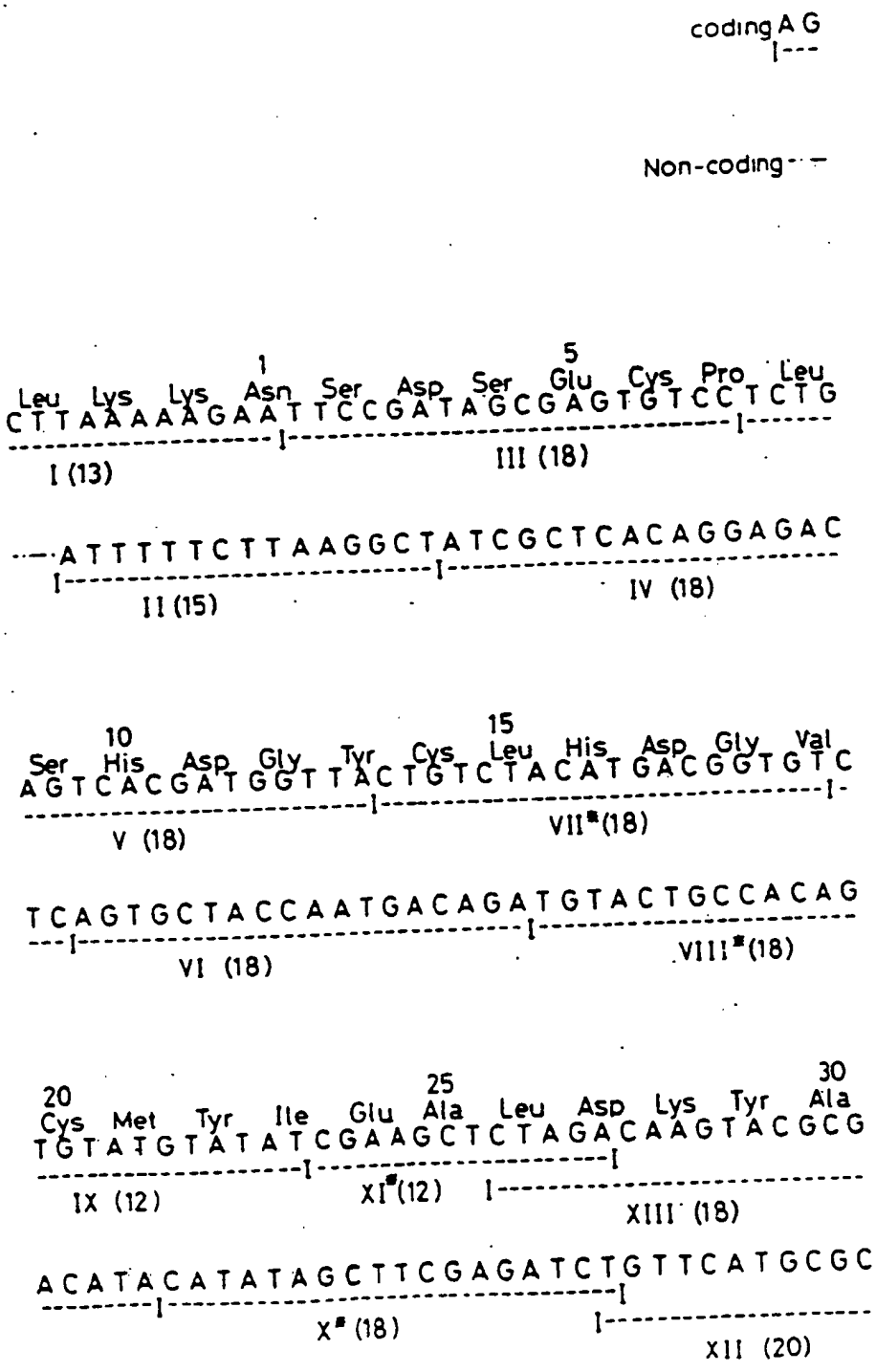


Fig. 4 (Part 1 of 2)

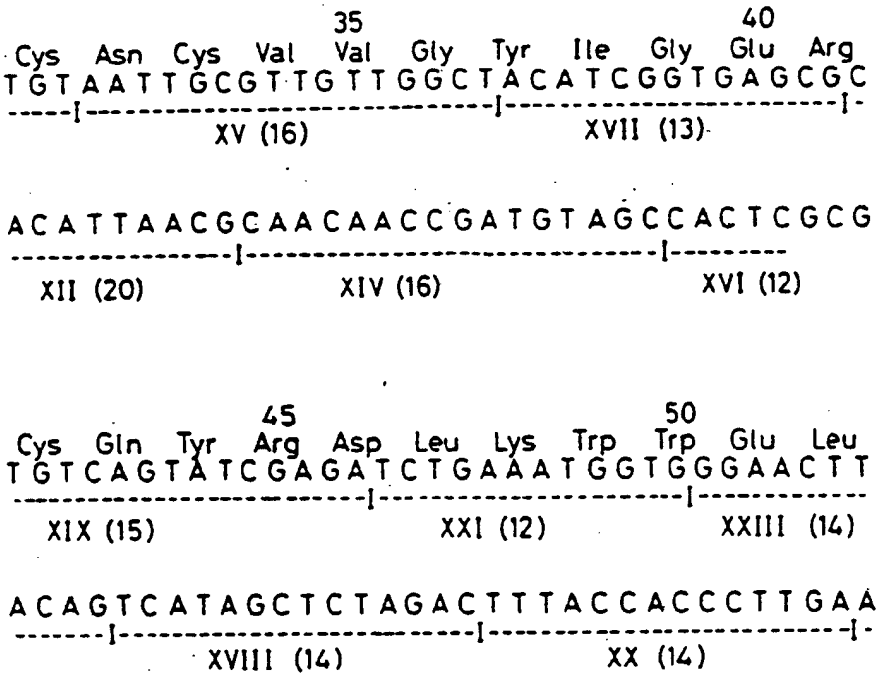


Fig. 4 (Part 2 of 2)

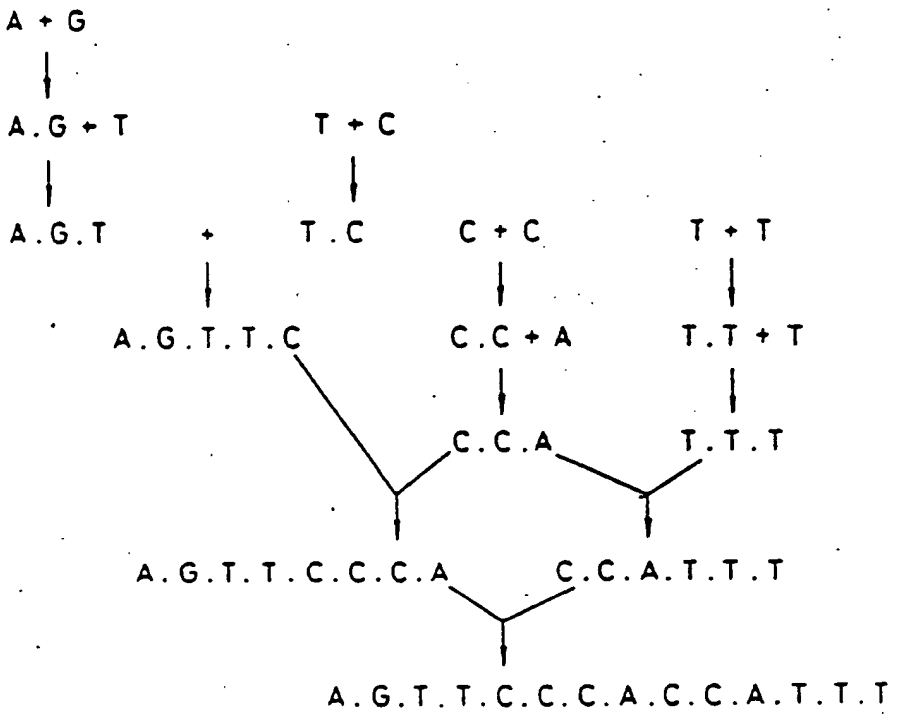

A + G
A.G + T
T + C
A.G.T
+
T.C
C + C
T + T
A.G.T.T.C
C.C + A
T.T + T
C.C.A
T.T.T
A.G.T.T.C.C.C.A
C.C.A.T.T.T
A.G.T.T.C.C.C.A.C.C.A.T.T.T

Fig.5

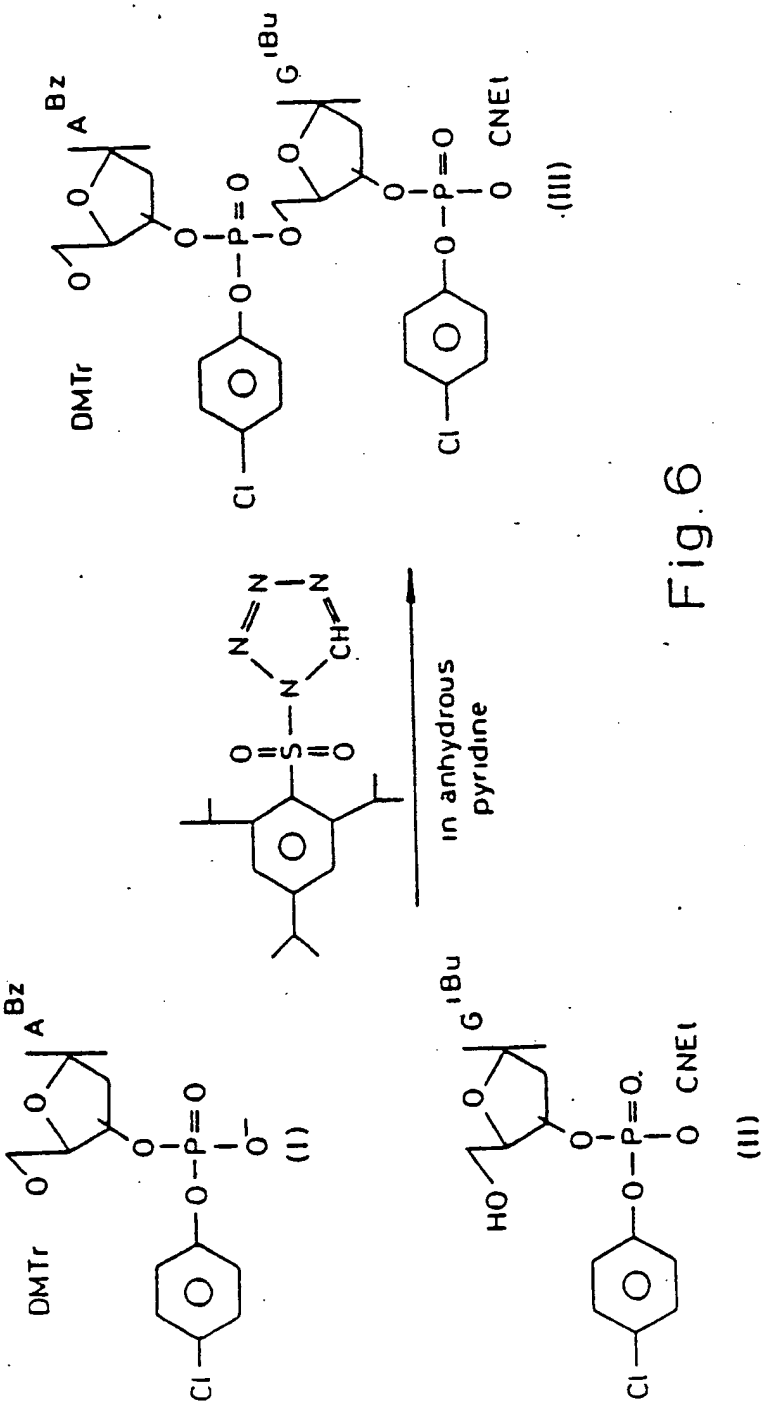


Fig.6

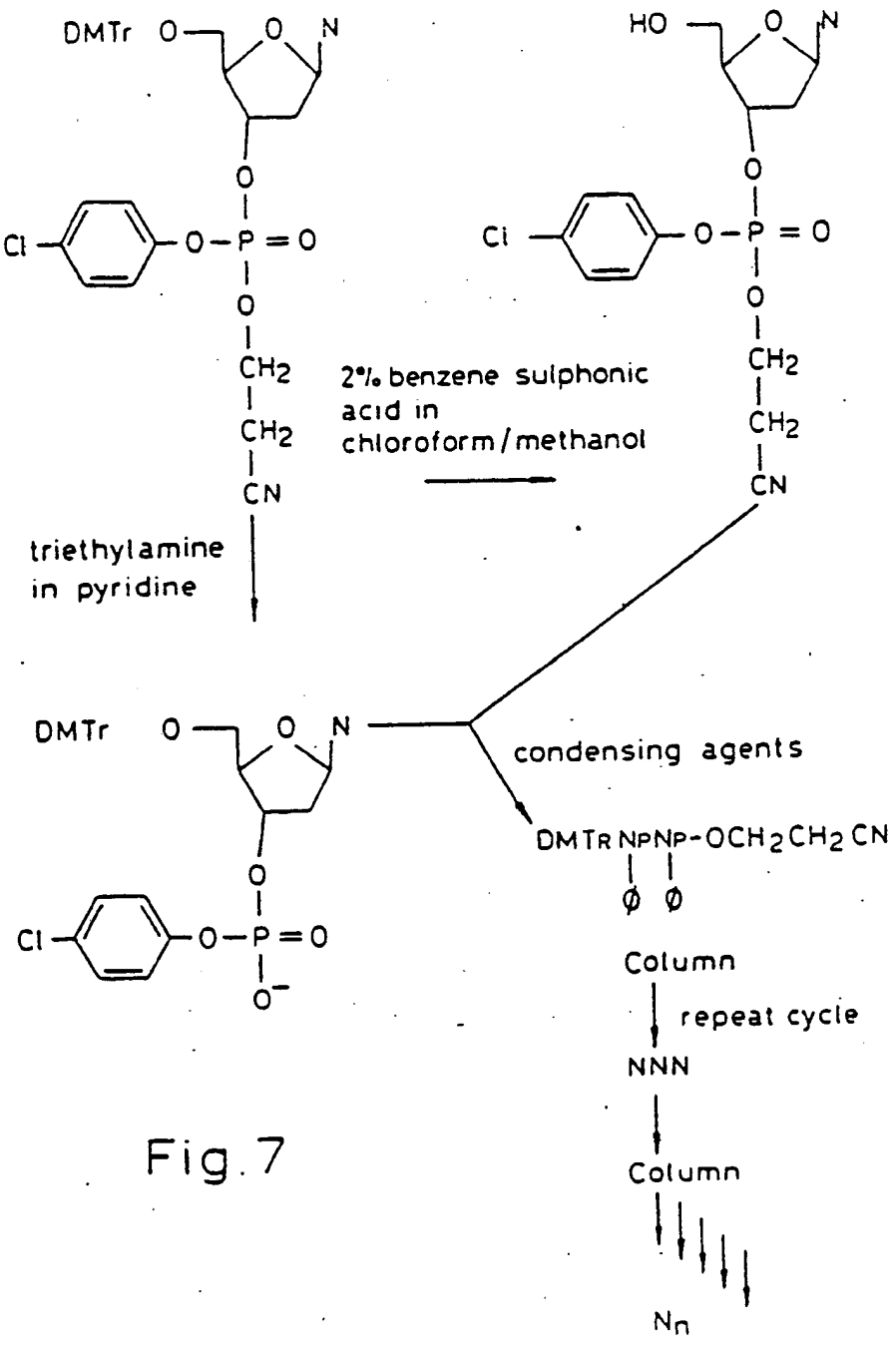

DMTr
HO
Cl
2% benzene sulphonic acid in chloroform/methanol
triethylamine in pyridine
condensing agents
DMTr NPNP-OCH2CH2CN
Column
repeat cycle
NNN
Column
Nn

Fig. 7

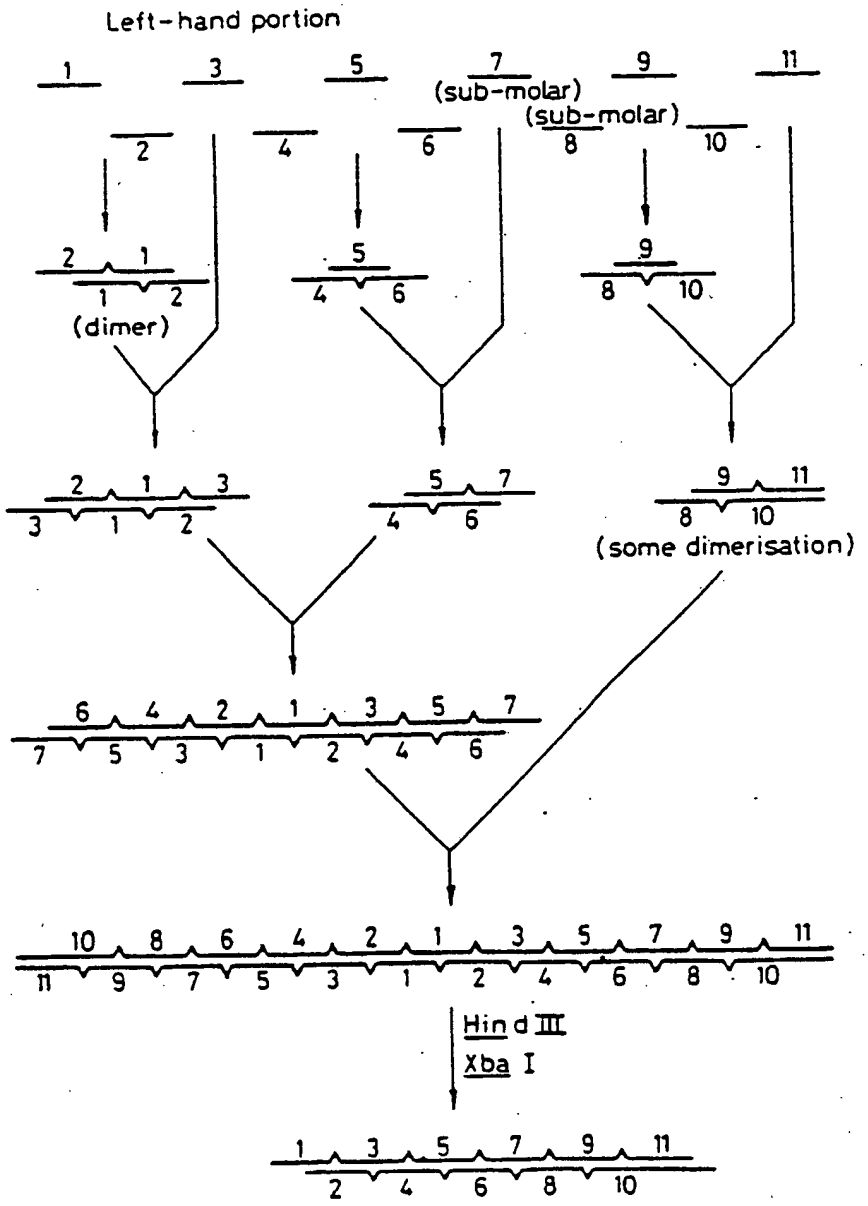


Fig.8 (Part 1 of 2)

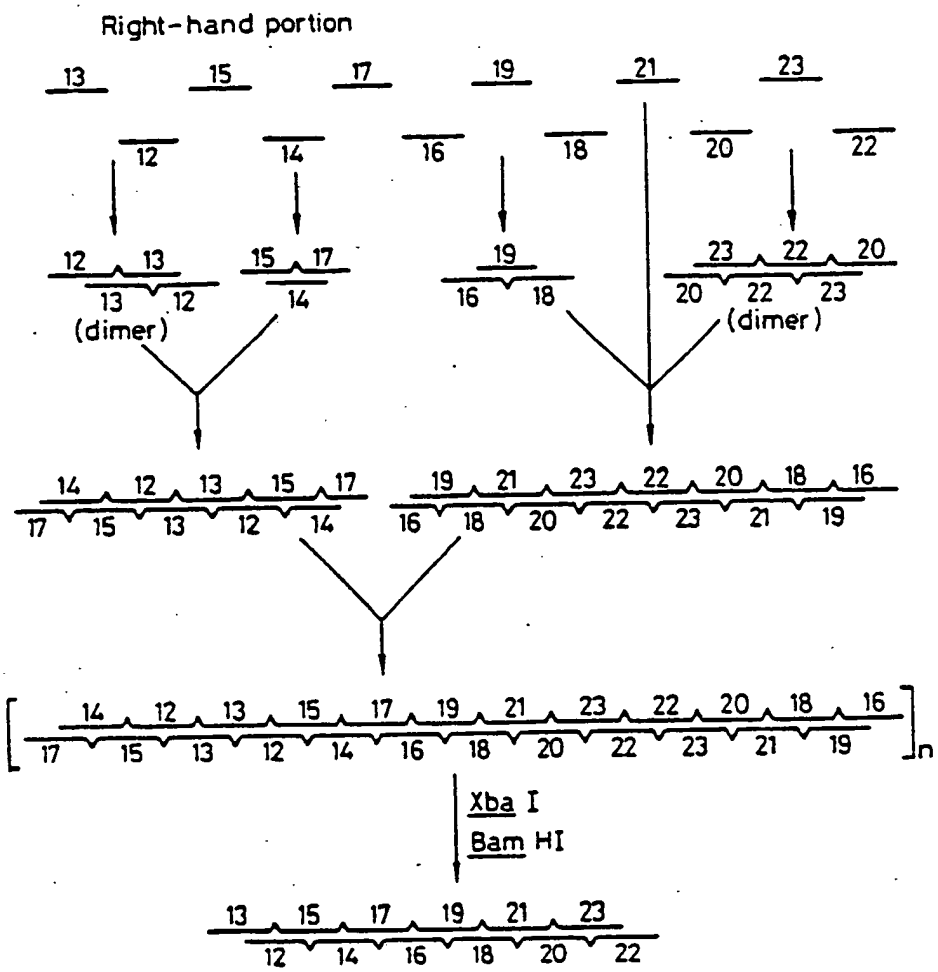


Fig.8 (Part 2 of 2)

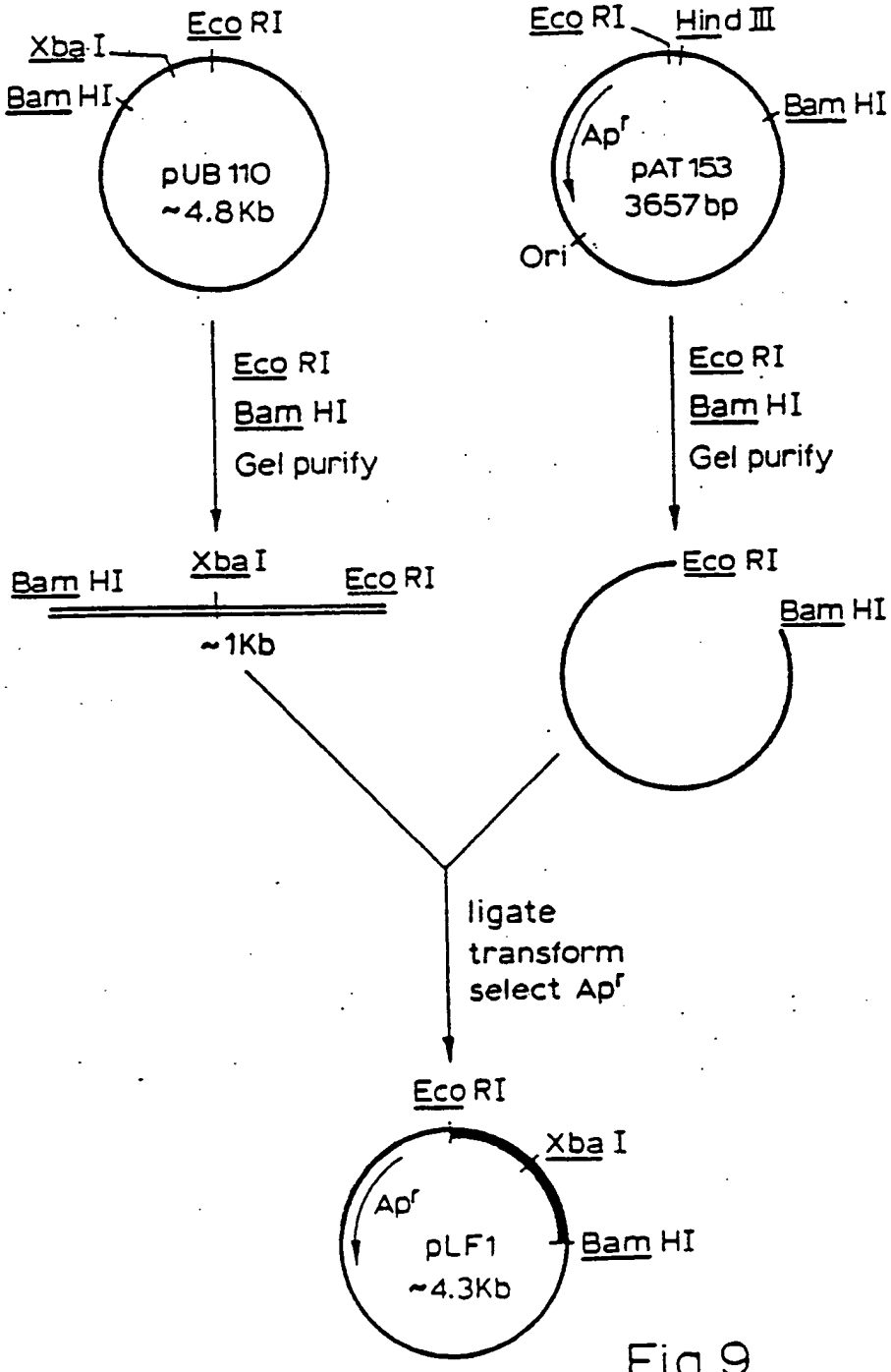

Eco RI
Xba I
Bam HI
pUB 110
~4.8 Kb
Eco RI
Hind III
Bam HI
Ap^r
pAT 153
3657 bp
Ori
Eco RI
Bam HI
Gel purify
Eco RI
Bam HI
Gel purify
Bam HI
Xba I
Eco RI
~1Kb
Eco RI
Bam HI
ligate
transform
select Ap^r
Eco RI
Xba I
Ap^r
pLF1
~4.3Kb
Bam HI

Fig.9

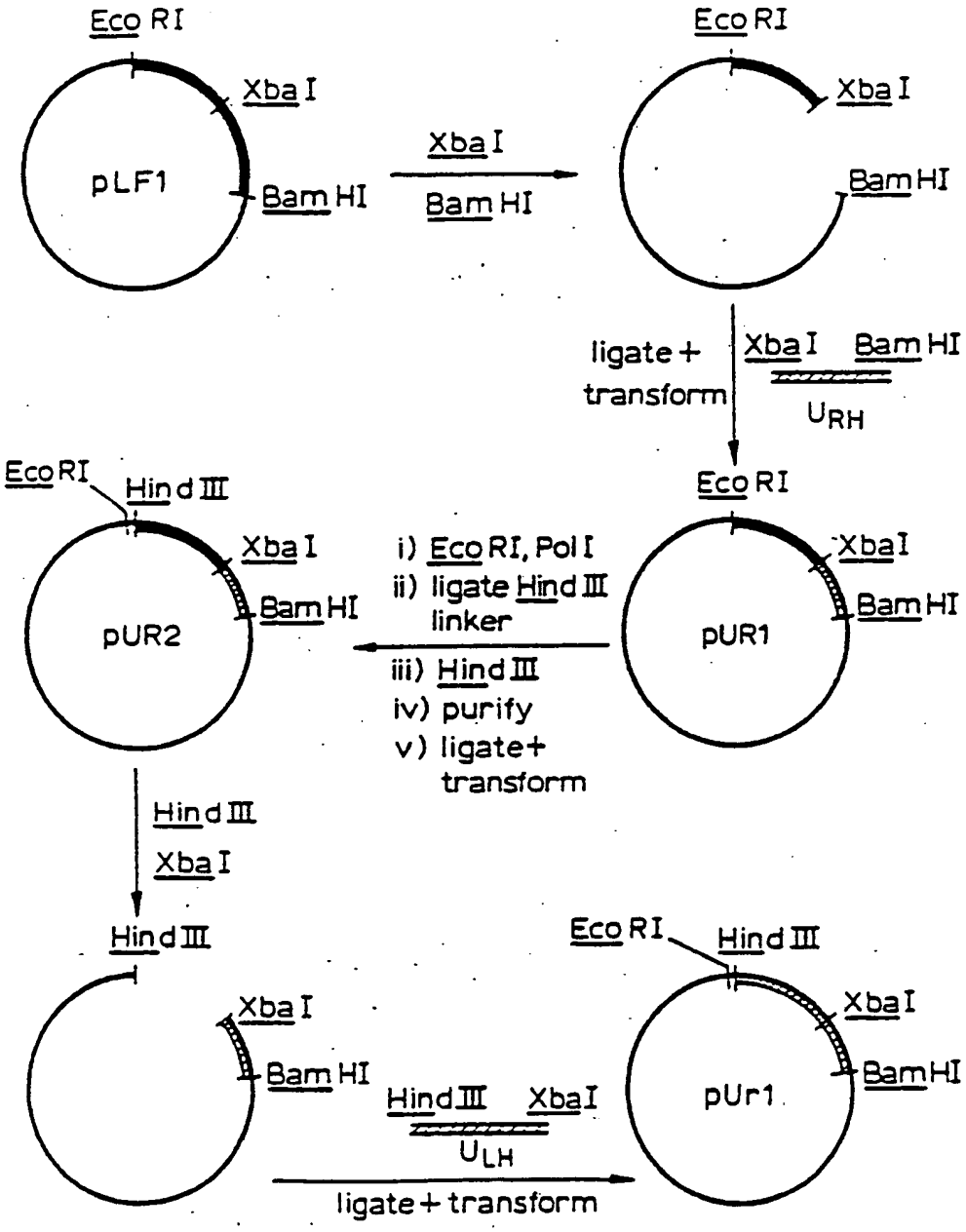

Eco RI
Xba I
pLF1
Bam HI
Xba I
Bam HI
Eco RI
Xba I
Bam HI
ligate + transform
Xba I Bam HI
URH
Eco RI
Hind III
Xba I
pUR2
Bam HI
i) Eco RI, Pol I
ii) ligate Hind III linker
iii) Hind III
iv) purify
v) ligate + transform
Eco RI
Xba I
pUR1
Bam HI
Hind III
Xba I
Hind III
Xba I
Bam HI
Hind III Xba I
ULH
ligate + transform
Eco RI
Hind III
Xba I
pUr1
Bam HI

Fig.10